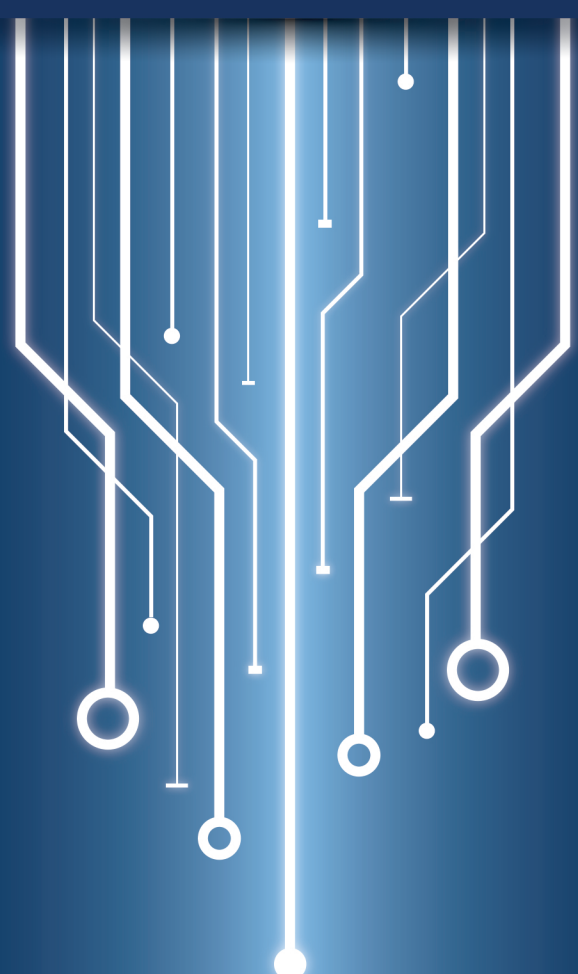
GET 800
CHOOSE YOUR COLLEGE

PHYSICS MASTERY

for the Advanced High School Student



Complete Physics Review with 400 SAT and AP Physics Questions

by Dr. Tony Rothman
Edited by Dr. Steve Warner

Legal Notice

This book is copyright 2016 with all rights reserved. It is illegal to copy, distribute, or create derivative works from this book in whole or in part or to contribute to the copying, distribution, or creating of derivative works of this book.

Physics Mastery for Advanced High School Students

Complete Physics Review with
400 SAT and AP Physics Questions

---

by Dr. Tony Rothman
Edited by: Dr. Steve Warner

© 2016, All Rights Reserved

BOOKS FROM THE GET 800 COLLECTION

28 New SAT Math Lessons to Improve Your Score in One Month
- Beginner Course
- Intermediate Course
- Advanced Course

320 SAT Math Problems arranged by Topic and Difficulty Level
New SAT Math Problems arranged by Topic and Difficulty Level
New SAT Verbal Prep Book for Reading and Writing Mastery
320 SAT Math Subject Test Problems arranged by Topic and Difficulty Level
- Level 1 Test
- Level 2 Test

The 32 Most Effective SAT Math Strategies
SAT Prep Official Study Guide Math Companion
Vocabulary Builder
320 ACT Math Problems arranged by Topic and Difficulty Level
320 GRE Math Problems arranged by Topic and Difficulty Level
320 SAT Math Problems arranged by Topic and Difficulty Level
320 AP Calculus AB Problems
320 AP Calculus BC Problems
SHSAT Verbal Prep Book to Improve Your Score in Two Months
555 Math IQ Questions for Middle School Students
555 Advanced Math Problems for Middle School Students
555 Geometry Problems for High School Students
Algebra Handbook for Gifted Middle School Students

CONNECT WITH DR. STEVE WARNER

Table of Contents

General Strategies for Solving Physics Problems 7

1. Physics Basics Review 11
    Problems Involving Dimensions, Units and Vectors 18

2. Kinematics Review 23
    Kinematics Problems 30

3. Particle Mechanics Review 38
    Particle Mechanics Problems 60

4. Rotation and Angular Momentum Review 86
    Rotation and Angular Momentum Problems 97

5. Circular Motion, Simple Harmonic Motion and Waves Review 103
    Circular Motion, Simple Harmonic Motion and Waves Problems 124

6. Reference Frames, Relative Motion and Relativity Review 137
    Relative Motion and Relativity Problems 146

7. Fluid Mechanics Review 150
    Fluid Mechanics Problems 158

8. Heat and Thermodynamics Review 165
    Thermodynamics Problems 182

9. Electricity and Magnetism Review 190
    Electricity and Magnetism Problems 210

10. Simple Circuits Review 230
    Simple Circuits Problems 241

11. Optics Review 252
    Optics Problems 276

12. Modern Physics Review **288**
Modern Physics Problems **306**

About the Author ***316***

Books from the Get 800 Collection ***317***

General Strategies for Solving Physics Problems

Physics Mastery has been designed to help you ace the SAT physics subject test and the AP physics 1 and AP physics 2 tests. To that end, the book is divided into review sections and problem sections. The review sections are, as their name implies, reviews. They are intended mostly to refresh your memory about material you should already know—at least have already seen—as well as to familiarize you with our notation and style. They do not replace a good text book (though they may replace a bad one) and the topics chosen are limited to those you will encounter on the College Board exams; they cannot and do not cover as much ground as a serious first-year physics course.

The last statement is a slight lie. In addition to reviewing the material needed to perform well on the subject tests and AP physics exams, we also present some material that is a bit more difficult than that which you will find on standardized tests. The hope is that mastering the more difficult material will make those tests easier for you.

For most people physics is a difficult subject, and nothing can replace actually understanding the material, which requires prolonged study and practice. Rote memorization at the eleventh hour is doomed to failure; memorization neither tells you which concepts are important and what equations to use, nor how to recognize a mistake. We hope that the problem-solving approach you will encounter here—which sometimes differs considerably from the methods taught in high school—will stay with you for the remainder of your careers. If you can do all the problems in *Physics Mastery*, you should perform very well on the College Board exams.

Broadly, we have attempted to keep the tone of the book conversational but concise. The SAT physics subject test and the AP physics 1 and AP physics 2 exams do not require calculus. Therefore, we do not either, although we have tried to be as rigorous as possible within that limitation, and have occasionally resorted to footnotes to explain concepts that really do require calculus for proper understanding.

The overall map of *Physics Mastery* is simple: easy to hard. In each chapter the review sections proceed from easier material to more difficult material. The chapters themselves proceed from elementary subjects to more demanding topics and the problems also proceed from basic subject test level questions to challenging AP-level problems. Admittedly, this scheme has been difficult to carry out in a consistent fashion because the division between subject test material and AP-level on the exams is not entirely clear and the same topics frequently appear on both exams. As a rule, however, subject test questions are more qualitative (conceptual), while the AP-level questions expect quantitative (mathematical) responses. The subject-level student is advised to read as much of the review material as possible, in particular the introductions to each section, which are mostly qualitative. To make life easier, each chapter is headed with a "warning label" as to which sections are at the subject test level and which are at the AP level. If you believe you have already mastered the material, you may proceed directly to the problems and consult the review sections as necessary.

The problems have been chosen to closely resemble those found on the exams, but we do not guarantee that you will find these exact problems on any exam. Most are standard exercises that you will encounter, not only on the College Board tests, but in any physics course. However, we give innumerable examples as well as solutions to every problem. Furthermore, for many problems we give multiple solutions: a quick solution, which may amount to an educated guess and which often involves **POE** (process of elimination), followed by a rigorous solution, sometimes an alternative solution, and occasionally **helpful** remarks that will undoubtedly prove to be just that. We have also found many of the College Board problems unclear or downright incorrect and have not shied away from fixing them. Throughout the review sections we also give hints, tips and advice, either boxed or in **boldface** that should make solving any problem easier. The most important strategies for problem solving we reveal right now:

Strategy 1: Draw a Diagram

As a rule, the first thing one should *always* do in solving a physics or math problem is make a good diagram. The SAT physics subject test usually provides these already, and so you are saved the trouble. If there is no diagram, draw one, even if it is just a vector pointing in the correct direction: ⟶

Strategy 2: Eliminate and Guess

SAT physics subject test questions are multiple choice and it pays to make a guess if you can eliminate most of the answers. If a problem has you going this way for 5 kilometers, that way for another 4, and a third way for 7, the final distance from your starting point cannot be 100 kilometers. If that is a choice, cross it off. Various ways you can eliminate answers are suggested in the other strategies below and in the problems.

Strategy 3: Avoid Excessive Calculator Use

Contrary to popular belief, the calculator is not a good way to eliminate answers. SAT subject test questions are conceptual and, at most, require only a few simple arithmetic operations. Calculators are superfluous (not to mention not allowed). Put yours away.

Because the questions are conceptual, logic is a useful weapon to eliminate incorrect choices. It may be that two choices are equally true or equally false, in which case neither can be the single correct response. Here is an example involving a box sliding down a frictionless ramp and hitting a spring (see problem 3.69).

Example 0.1: The maximum velocity of the box occurs at the instant when

(A) the box hits the spring.
(B) the potential energy is minimum.
(C) the kinetic energy is maximum.
(D) the forces on the block are zero.
(E) all of the above except (A)

Even without knowing exactly what the problem is about, you can arrive at the correct answer. By definition, maximum velocity implies that kinetic energy is a maximum. But because energy is conserved, (B) and (C) mean exactly the same thing. Therefore, there are at least two correct responses to the question and, logically, the only possible answer to the question is (E).

Strategy 4: Dimensional Analysis

In terms of physics itself, the most powerful weapon for checking answers (and one that you should use routinely in *every* step of *every* calculation you do from now until the day you die) is what physicists term **dimensional analysis**. Section 1.2, and a few of the problems in Chapter 1 give more detail on this crucial technique, but here is the basic idea:

Any quantity in physics comes with a set of dimensions—how it is expressed in terms of the fundamental physical quantities of mass, length and time.

Example 0.2: A spring with spring constant k is attached to a mass m that is oscillating in Earth's gravitational field, which produces an acceleration g. Which of the following expressions might be the spring's period?

(A) k/m
(B) k/gm
(C) $\sqrt{m/k}$
(D) $\sqrt{g^2/km}$
(E) (B) or (D)

You might remember the answer, but let's say you don't. You can decide from the basics. First, you tell yourself that the period of anything is a time, t, so whatever your answer is, it had better have dimensions of t, or it doesn't have a chance of being correct. You might remember Hooke's law (Section 3.5), which says that the force exerted by a spring is $F = -kx$. You certainly remember $F = ma$, and that an acceleration has the dimensions of length/(time^2) $= \ell/t^2$. Thus, force has the dimensions of $m\ell/t^2$. Since x is also a length ℓ, dimensionally we have $k\ell = m\ell/t^2$. Hence, k/m has dimensions of $1/t^2$ and therefore m/k has dimensions of t^2. We want t, and so the correct answer must be (C).

You might wonder whether any of the other answers can be correct. **Hint: Whenever a question asks, "given two (or three) quantities, make a quantity with dimensions of such-and-such," there is always a single, unique answer.** (C) is the only possibility. The hint also means it is impossible for both (B) and (D) to be correct at the same time, ruling out (E) immediately.

Strategy 5: Think About the Numbers

A second powerful tool in eliminating incorrect physical choices is a number sense. It is useful to create a **number bank** with which you can perform a **reality check** on your answers.

Example 0.3: A teacher is recording her mass with a precision of one-half percent. What is the most likely result?

(A) 6.43 kg
(B) 60 kg
(C) 64.3 kg
(D) 600 kg
(E) 643 kg

Now, only two of these candidates lie within a plausible mass range for an adult human. If you have a sense of what a realistic mass for a person is, you can eliminate all but choices (B) and (C) at once. Similarly, it is useful to have at your fingertips a sense about important sizes (continents, the Earth, the distance to the Moon, the diameter of the Sun, the size of an atom, and so on) as well as a sense of masses of the various particles that appear in physics (atoms, electrons, protons, people, etc.).

Strategy 6: Estimate Answers

Having a good number sense goes hand in hand with being able to make an estimate.

Example 0.4: Calculate the Earth's mean density given that the Earth's gravitational acceleration is $g = 9.8\ \text{m/s}^2$, its equatorial radius is $R = 6.378 \times 10^6\ \text{m}$ and the gravitational constant is $G = 6.67 \times 10^{-11}\ m^3/(\text{kg s}^2)$.

You might say that the Earth is a sphere with mass $M = \rho V$, density ρ and volume V. Then

$$g = \frac{GM}{R^2} = \frac{G\rho V}{R^2} = \frac{G\rho\left(\frac{4}{3}\pi R^3\right)}{R^2} = G\rho\left(\frac{4}{3}\pi\right)R$$

At that point you could plug all your numbers into your calculator and find that

$$\rho \approx 5500\ \text{kg/m}^3 = 5.5\ \text{gm/cm}^3.$$

On the other hand, $g \approx 10\ \text{m/s}^2$, where $\approx$ means "approximately equal to." Of course $\pi \approx 3$. Then

$$\rho \approx \frac{g}{4GR} \approx \frac{10}{4\times 40\times 10^{-5}} = \frac{10^6}{160} = 6250\ \text{kg/m}^3 \approx 6\ \text{gm/cm}^3.$$

For all the effort in putting in all those decimal points, the calculator's answer is hardly better than a quick estimate. *This density is a good number to put in your number bank*: Most of the things we encounter in daily life—water, wood, dirt, rocks, metal—have a density in the range $1 - 10\ \mathrm{gm/cm^3}$, and if you got an answer that was ten times larger or smaller, a reality check would tell you that something went wrong.

AP problems involve more elaborate calculations than SAT subject test problems, but you should nevertheless be able to use your number sense to approximate the quantities involved, round off numbers and arrive at an answer that is within a factor of ten or better than one that you'd get with your calculator. We call this making an order-of-magnitude estimate (traditionally a "back-of-the-envelope calculation"). If your estimate and your numerical answer do not agree, you should rework the problem until they do. When you are finished, you should always subject your answer to a reality check; if the answer does not accord with your sense of reality, redo the problem. On an exam you probably do not have time to make careful estimates, but always perform a brief mental reality check to make sure the answer is plausible.

Strategy 7: Use Proportions

Often SAT and AP problems ask, "by what factor does such-and-such a quantity change if you do so-and-so." That wording is exam code to use proportions. The following problem actually appeared on an exam at Princeton:

Example 0.5: Transmitted electrical power is $P_t = IV$, where I is the current and V is the voltage. Most electricity to our homes is transmitted at about 100,000 volts. On the other hand, wires dissipate electrical power as heat according to $P_d = I^2R$, where R is the resistance of the wire. Resistance in turn is given by $R = \rho L/A$, where ρ is a property of the material called resistivity, L is the length of the wire and A is the wire's cross-sectional area.

The resistance per kilometer of the wire is, say, 1 ohm and the radius of the wire is 1 cm. If the voltage is lowered from 100,000 to 100 volts and you want to keep the transmitted power P_t and P_d the same as they were previously, how much does the weight of the wire change?

The vast majority of the students used their calculators to compute the area, compute the resistance, find a new current and then recalculate everything and get a numerical answer that was wrong. All that is required is proportions: If the voltage drops from 100,000 to 100 V, a factor of 1000, to keep P_t constant requires that I go up by the same factor of 1000. $P_d = I^2R$, and so to keep P_d constant you must then lower R by a factor of 1 million. For constant ρ, the resistance per kilometer is proportional to $1/A$; thus, you must increase the area by a factor of 1 million. The weight per kilometer of the wire is proportional to the area, and so the weight also goes up by a factor of one million. This is why we transmit electrical power at high voltage.

Never work out a problem twice if you can avoid it. Form a ratio of the required quantities and check the factor by which the answer changes.

All of these strategies point to the **Prime Directive** in physics:

Never Do a Calculation Until You Already Know the Answer

It takes a lot of practice to be able to see an answer before you begin calculations, but to give you that practice is the purpose of this book.

1. PHYSICS BASICS REVIEW

Sections 1-4 of this chapter are at the subject-test level. Sections 5-7 are at the AP level. However, use of the right-hand-rule from Section 7 is required for the subject test.

1. Essential Terms

The word **physics** is from the Greek, *physika* meaning "natural things," and physics is the study of the material universe. Physics is the most basic science. That is why you love it.

All quantities in physics come with a set of **dimensions**, which tell you how it is expressed in terms of the fundamental quantities of **mass**, **length**, and **time**. Every dimension can be expressed in many different **units**. For example, mass can be measured in grams or kilograms. Most introductory texts, including this one, use the **Système Internationale**, otherwise known as the **MKS** (meter-kilogram-second) system, although most physicists avoid it.

All quantities in elementary physics can also be classified into two groups: **scalars** and **vectors**. Scalars are entirely characterized by their size, or magnitude (speed, temperature), while vectors are characterized by their **magnitude** and **direction** (velocity, acceleration).

Few SAT and AP problems ask direct questions about dimensions, units, scalars and vectors; however, the exams do expect you to recognize the meaning of the units when you encounter them. They also underlie the remaining material in this book, so we review them.

2. Dimensions and Units

All quantities in an elementary physics course and on an SAT or AP exam can be derived from five fundamental quantities: mass (M), length (L), time (T), charge (Q) and temperature (Θ). The last two appear only in problems involving electricity and thermodynamics. These five fundamental quantities are termed **dimensions**, in analogy to length, width and height. For example, since velocity is measured as a distance/time, we say the dimensions of velocity are length divided by time. Often this is written $[v] = \mathrm{L/T}$ or $[v] = \mathrm{LT}^{-1}$, the square brackets denoting dimensions.[1] In terms of problem solving, absorb the

Supreme Hint:

The dimensions of your answer must be correct. If the dimensions are wrong, the answer cannot be right.

If the numerical value of your answer is incorrect, this is a mistake. If the dimensions are wrong, the answer is meaningless. It is not even wrong. *You can often rule out answers on standardized tests by checking dimensions.* Some answers may have incorrect dimensions.

---

[1] Although the use of upper-case letters for dimensions has become the "official" method, it can be confusing, not least because T is also used for temperature and writing m for a mass but M for dimensions of mass in the same line does not elucidate much. Most physicists in fact use lower case m, ℓ and t for the dimensions of mass, length and time, as we sometimes will, unofficially. See problem 7 below for a comparison of both methods. We also speak of the "dimensions of force" or "dimensions of energy," meaning quantities that have the dimensions of force or energy, even though these are not one of the five fundamental dimensions.

Although the dimensions of velocity are always length/time, there are any number of different **units** in which the dimensions can be measured. For instance, length can be measured in feet, miles, kilometers, centimeters, meters, furlongs… Time can be measured in hours, days, years, seconds, blinks of an eye…

Semi-Supreme Hint:

The units of your answer must be consistent. If they are not, you cannot get the correct numerical answer.

Although most physicists hate it, introductory texts and SAT exams use the metric "Système Internationale," or SI, units. When only distance, mass and time appear, this is the "MKS," or meter-kilogram-second, system.

Dimension	Unit
length	meter (m)
mass	kilogram (kg)
time	second (s)

All basic physics involves forming quantities from these fundamental ones, determining whether they are important and learning how they behave.

Dimensions	Units	Quantity
L/T	m/s	speed, velocity
L/T^2	m/s^2	acceleration
ML/T	kg m/s	momentum
ML/T^2	$kg\ m/s^2$	force
ML^2/T^2	$kg\ m^2/s^2$	energy

In speaking of units and dimensions we employ the equivalence symbol: ≡. For example, we might write

$$1\ \mathrm{N} \equiv 1\ \mathrm{kg\ m/s^2},$$

which reads "1 newton is defined as 1 kilogram-meter per second squared." The equivalence sign merely signifies that this is a definition, so you can stop worrying about why it is correct. It is also expected that you know how to convert from the CGS system (centimeter-gram-second) to the MKS system.

All the above quantities are extremely important. **Memorize them!**

Example 1.1: Units conversion. A furlong is an obsolete English system unit equal to 201.2 m. Convert furlongs per fortnight to meters per second.

A fortnight is two weeks. Thus, 1 fortnight equals $2 \times 7 \times 24 \times 60 \times 60 = 1.2 \times 10^6$ s. (**Exercise**: convince yourself of this number.). A furlong per fortnight is therefore $202.2/(1.2 \times 10^6) = 1.67 \times 10^{-4}$ m/s, or 0.167 mm/s, slower than a snail's pace, perhaps a molasses's drip.

Note that one day is a little less than 10^5 s, a good number for your number bank.

3. Scalars

All quantities in elementary physics naturally fall into one of two categories that have very different properties: **scalars** and **vectors**. Because they have different properties, they must be treated differently.

A **scalar** is a quantity that can be described by specifying a single number: its **magnitude**, or size. Common examples of scalars encountered in College Board problems are: distance, speed, mass, temperature, density, height, energy, time. Once you have given a scalar's magnitude, you have said all there is to say about it. "The car is moving at 90 kilometers per hour," or "The car is moving at 10 meters per second," fully specifies the speed of the car. Scalars are usually displayed in italics: $v = 90$ km/h.

Scalar and vector addition crop up everywhere in physics and so it is important to understand how to do this properly. For scalars it is easy.

Scalar addition: Two scalars are added by adding their magnitude (there is nothing else to add). The total height above ground of a 2-meter-tall man standing on a 50-meter-tall building is 52 meters. A 5 kg mass added to a 6 kg mass gives an 11 kg mass.

4. Vectors

A **vector** is a quantity that must be described by specifying two properties: a **magnitude** and a **direction**. Common examples of vectors are: displacement, velocity, acceleration, momentum, force. "The car is moving 60 miles per hour at a heading of 90°," gives the car's velocity, as opposed to speed, which is the magnitude of the velocity. The best, and most common, way to visualize a vector is as an arrow:

The arrow's length gives the vector's magnitude and the tip points in the vector's direction. Vectors are generally displayed by either bold face type, such as $\mathbf{v}$, or by putting an arrow over the quantity: $\vec{v}$. We will use the former throughout this book. The magnitude, or length. of a vector $\mathbf{v}$ is designated by $|\mathbf{v}|$ or just v. For the above car, we could write $v = |\mathbf{v}| = 60$ mph. *The magnitude of a vector is a scalar*.

Note: When you move a vector parallel to itself, neither its length nor direction changes. Thus, it remains the same vector.

Warning: the length of a vector is usually *not* physical length; its length has the dimensions and units of the quantity it represents, for example, mph or newtons.

Every time you add two forces or momenta together, you are adding vectors. Indeed, these quantities naturally have a direction. This is why knowing how to add vectors is crucial. The rules for vector addition are much more complicated than for scalar addition. The first thing to remember is that scalars can be added only to scalars and vectors can be added only to vectors. You cannot add a scalar to a vector. They are apples and oranges.

Vector addition: When adding vectors, their directions must be taken into account. Vectors in opposite directions tend to "cancel each other out."

Tip: Two vectors of equal magnitude but opposite direction sum to zero.

In general, two vectors $\mathbf{A}$ and $\mathbf{B}$ are added by placing them "head to tail" as in Figure 1.1. The **resultant vector** $\mathbf{C} = \mathbf{A} + \mathbf{B}$ is by definition the vector drawn from the tail of $\mathbf{A}$ to the tip of $\mathbf{B}$.

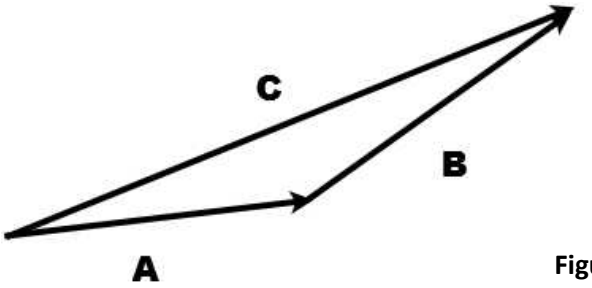


Figure 1.1

The negative of a vector is given by rotating it 180°. Also, $\mathbf{C} = \mathbf{A} - \mathbf{B}$ is by definition given by

$$\mathbf{C} = \mathbf{A} + (-\mathbf{B}):$$

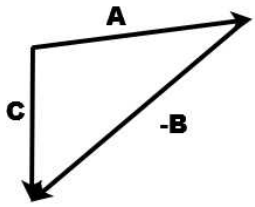


Figure 1.2

One almost always wants to know the magnitude of the resultant, $|\mathbf{C}|$. To find $|\mathbf{C}|$ you can draw it, as above, and measure its length. This gives an approximate answer. To find the exact magnitude of a resultant, the general prescription is:

1. Find the x- and y-components of **A** and **B**. ("x-component" refers to the part of the vector lying in the x-direction. In three dimensions, there are x-, y- and z-components.)
2. Add the x-components of **A** to the x-components of **B**. Similarly, for the y-components (and z-components if there is a third dimension).
3. Square the sums and add them.
4. Take the square root.

Essentially, **one uses the Pythagorean theorem to find the magnitude of the resultant.**

Example 1.2: Adding vectors.

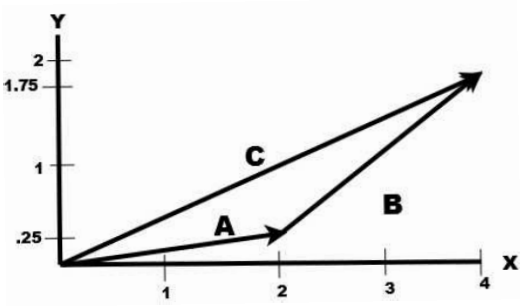


Figure 1.3

Here, the vector **A** has x-component 2 and y-component 0.25. Vector **B** has x-component 2 and y-component 1.5. (Why? Think parallel.) Thus, vector **C** has x-component $2 + 2 = 4$ and y-component $.25 + 1.5 = 1.75$. The magnitude of **C** is $|\mathbf{C}| = \sqrt{4^2 + 1.75^2} = 4.37$.

This example illustrates several important properties of vectors:

The magnitude of $\mathbf{C} = \mathbf{A} + \mathbf{B}$ is generally NOT the magnitude of A plus the magnitude of B.

You must always add x-components to x-components and y-components to y-components.

The maximum possible magnitude of $\mathbf{C} = \mathbf{A} + \mathbf{B}$ occurs when **A** and **B** are pointing in the same direction. Then $|\mathbf{C}| = |\mathbf{A}| + |\mathbf{B}|$. The minimum possible magnitude of **C** will be when **A** and **B** are pointed in opposite directions. Then $|\mathbf{C}| = |\mathbf{A} - \mathbf{B}|$.

5. Unit Vectors

As the name implies, a **unit vector** is a vector with magnitude 1. Unit vectors are important for several reasons. In the first place, since they have unit magnitude, they can be considered "pure directions" or "direction pointers." Furthermore, any arbitrary vector can be decomposed into unit vectors, in particular, the unit vectors lying along the x-, y- and z-axes. The x-, y- and z- unit vectors are usually written $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$ (read "i-hat, j-hat and k-hat") and are illustrated below.

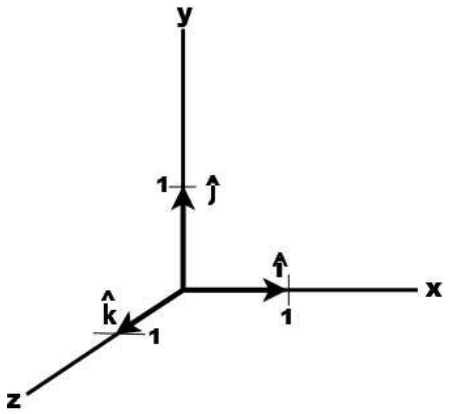


Figure 1.4

A vector of length 6 lying along the x-axis may thus be written as $\mathbf{A} = 6\hat{\mathbf{i}}$**,** while a vector of length 7 lying along the y-axis may be written as $\mathbf{B} = 7\hat{\mathbf{j}}$. The 6 and 7 represent the magnitudes and $\hat{\mathbf{i}}$ and $\hat{\mathbf{j}}$ give the directions**.** In Example 1.2 above, we could have written $\mathbf{A} = 2\hat{\mathbf{i}} + 0.25\hat{\mathbf{j}}$ and $\mathbf{B} = 2\hat{\mathbf{i}} + 1.5\hat{\mathbf{j}}$. In doing this we have *resolved* **A** and **B** into x- and y-components. Then **C** is obtained by adding the components separately: $\mathbf{C} = 4\hat{\mathbf{i}} + 1.75\hat{\mathbf{j}}$, and $|\mathbf{C}|$ is found by squaring the components separately, adding them and taking the square root: $|\mathbf{C}| = \sqrt{4^2 + 1.75^2} = 4.37$, as we already determined.

Whenever you are presented with a vector, you should *immediately* think of resolving it into components. This is especially necessary in force and momentum problems.

Given any vector in any direction, one can always convert it into a unit vector by dividing the vector by its magnitude.

Example 1.3: Forming unit vectors. Let $\mathbf{A} = 2\hat{\mathbf{i}} + 3\hat{\mathbf{j}}$. Then the magnitude of **A** is $\sqrt{13}$. Therefore, $\hat{\mathbf{A}} = \frac{1}{\sqrt{13}}(2\hat{\mathbf{i}} + 3\hat{\mathbf{j}})$ is a unit vector pointing in the direction of **A**.

6. Scalar Products

Just as one can add vectors, one can also multiply them. However, while it is clear that $3 \times 5 = 15$, it is not so clear how to multiply two vectors **A** and **B** when they both have not only a magnitude (3 and 5) but a direction as well (north and north west). In elementary physics, two types of vector multiplication prove extremely useful.

The **scalar product**, or **dot product**, of two vectors **A** and **B** is defined as

$$\mathbf{A} \cdot \mathbf{B} \equiv AB\cos\theta \quad (1.1)$$

where A and B are the magnitudes of **A** and **B**, and θ is the angle between them (Figure 1.5). It is called the scalar product because, as you see from the definition, the result is a scalar rather than a vector.

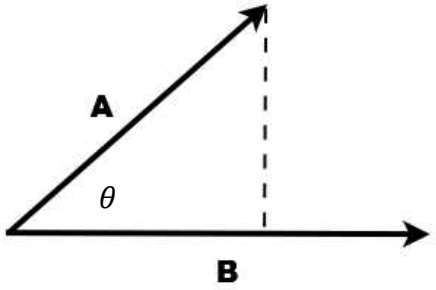


Figure 1.5

Note several important properties of the dot product. First, from the definition you can see that the dot product is *commutative*. In other words, $\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A}$. You should also check that the dot product is *distributive* over addition. That is, $\mathbf{A} \cdot (\mathbf{B} + \mathbf{C}) = \mathbf{A} \cdot \mathbf{B} + \mathbf{A} \cdot \mathbf{C}$.

Second, the definition implies that for any two parallel vectors, $\mathbf{A} \cdot \mathbf{B} = AB\cos(0) = AB$. In other words, $\mathbf{A} \cdot \mathbf{B}$ becomes just the product of their lengths. In particular, $\mathbf{A} \cdot \mathbf{A} = A^2$. This last fact is extremely valuable: **dotting a vector into itself gives the square of the magnitude.**

The previous statement applies equally to unit vectors. That is, $\hat{\mathbf{i}} \cdot \hat{\mathbf{i}} = |\hat{\mathbf{i}}|^2 = 1 = \hat{\mathbf{j}} \cdot \hat{\mathbf{j}} = \hat{\mathbf{k}} \cdot \hat{\mathbf{k}}$. Similarly, because $\hat{\mathbf{i}}, \hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$ are all at right angles to one another, $\theta = 90°$, which implies that $\hat{\mathbf{i}} \cdot \hat{\mathbf{j}} = \hat{\mathbf{i}} \cdot \hat{\mathbf{k}} = \hat{\mathbf{j}} \cdot \hat{\mathbf{k}} = 0$.

Let us write two arbitrary vectors as $\mathbf{A} = A_x\,\hat{\mathbf{i}} + A_y\,\hat{\mathbf{j}}$ and $\mathbf{B} = B_x\,\hat{\mathbf{i}} + B_y\,\hat{\mathbf{j}}$, where the subscripts x and y refer to the x- and y-components of the vectors (see Example 1.2 above). Then, taking the dot product, gives

$$\begin{aligned}\mathbf{A} \cdot \mathbf{B} = \ & (A_x\,\hat{\mathbf{i}} + A_y\,\hat{\mathbf{j}}) \cdot (B_x\,\hat{\mathbf{i}} + B_y\,\hat{\mathbf{j}}) = A_xB_x\,\hat{\mathbf{i}} \cdot \hat{\mathbf{i}} + A_xB_y\,\hat{\mathbf{i}} \cdot \hat{\mathbf{j}} + A_yB_x\,\hat{\mathbf{j}} \cdot \hat{\mathbf{i}} + A_yB_y\,\hat{\mathbf{j}} \cdot \hat{\mathbf{j}} \\ & = A_xB_x + A_yB_y\end{aligned}$$

This is often regarded as an alternative definition of the dot product:

$$\mathbf{A} \cdot \mathbf{B} = A_xB_x + A_yB_y \quad (1.2)$$

Note that $\mathbf{A} \cdot \mathbf{A} = A_x^2 + A_y^2 = A^2$ by the Pythagorean theorem. (**Exercise:** Show that definition (1.2) is equivalent to definition (1.1)).

Finally, referring to Figure 1.5, because $A\cos\theta$ is the length of **A** that lies along **B**, we can also write

$$\mathbf{A} \cdot \mathbf{B} = B \times (\text{the projection of } A \text{ on } B).$$

Also convince yourself that

$$\mathbf{A} \cdot \mathbf{B} = A \times (\text{the projection of } B \text{ on } A).$$

Note that in the above two formulas, the symbol $\times$ is being used for ordinary multiplication of scalars. This shouldn't be confused with the cross product, a multiplication between two vectors, which we define below.

Example 1.4: Angles between vectors. One of the most common applications of the dot product is to find the angles between vectors. From definition Eq. (1.1), as long as **A** and **B** are not the zero vector,

$$\frac{\mathbf{A}}{A}\cdot\frac{\mathbf{B}}{B}=cos\theta\ .$$

But $\mathbf{A}/A$ is by definition $\hat{\mathbf{A}}$, the unit vector in the direction of **A**, with a similar expression for $\hat{\mathbf{B}}$. Thus, to find the angle between two vectors **A** and **B**, you merely need to form the unit vectors in their direction, take the dot product and then the inverse cosine. If $\mathbf{A}=2\hat{\mathbf{i}}+4\hat{\mathbf{j}}$ and $\mathbf{B}=3\hat{\mathbf{i}}-2\hat{\mathbf{j}}$, we have $A=2\sqrt{5}$ and $B=\sqrt{13}$. Then

$$\hat{\mathbf{A}}\cdot\hat{\mathbf{B}}=\frac{1}{2\sqrt{65}}(6-8)=\cos\theta,$$

or $\cos\theta=-0.124$. Taking the inverse cosine gives $\theta=97.1$. (**Exercise:** Accurately draw the vectors and confirm the calculation visually.)

7. Vector Products

The **vector product**, or **cross product**, of two vectors **A** and **B** is defined, in contrast to the dot product, as a vector **C**. The standard notation is $\mathbf{C}=\mathbf{A}\times\mathbf{B}$ (read "C = A cross B"). If **C** is the cross product of **A** and **B**, then

$$C\equiv AB\sin\theta \quad (1.3)$$

C points in the direction given by the right-hand rule.

Here, only the *magnitude* of $\mathbf{C}=C$ is given by $AB\sin\theta$, where, again, θ is the angle between **A** and **B**.

The direction of **C** is given by the infamous **right-hand rule (RHR)**, which says: To cross **A** into **B**, point the fingers of your right hand along the direction of **A**. Curl your fingers in the direction of **B**. Your thumb will point in the direction of **C** (Figure 1.6). Alternatively, you can visualize a screw with a normal right-hand thread sitting at the junction of **A** and **B**. If you imagine inserting a screwdriver into the screw and turning it so that **A** rotates into **B,** the screw will advance in the direction of **C**.

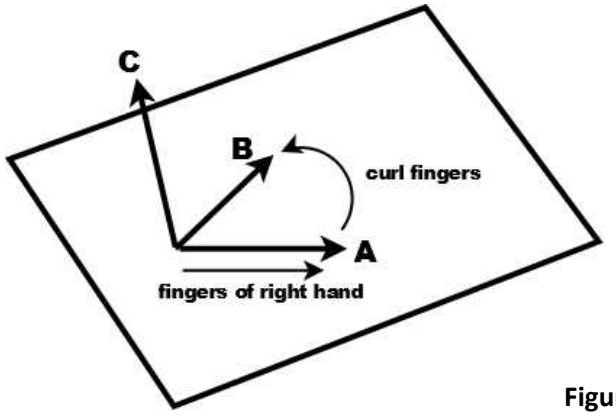


Figure 1.6

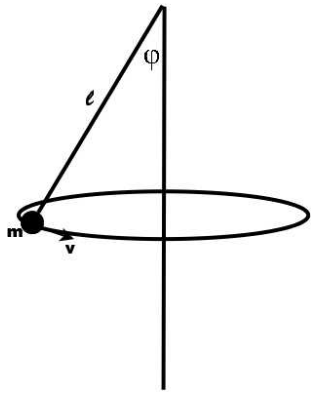
Note: C is at right angles to both A and B.

Warning: To get the right answer by the right-hand rule, you must use your right hand.

The warning is another way of stating that $\mathbf{A}\times\mathbf{B}=-\mathbf{B}\times\mathbf{A}$. *The cross product is not commutative*.

Although the cross product initially seems strange, it turns out to be extremely useful because in certain situations involving torques and magnetic fields (see Sections 4.6 and 9.10) a force may act at right angles to the direction of an object's motion.

Example 1.5: Ball on string. A ball is attached to the ceiling and is whirling around on a string, as shown. If $\boldsymbol{\ell}$ is the vector pointing from the point of attachment to the ball, and **v** is the vector pointing in the direction of the ball's motion (the velocity), in what direction does $\mathbf{v} \times \boldsymbol{\ell}$ point?



Using the RHR, the resultant points up and to the left, at an angle φ from the horizontal. (**Exercise:** Convince yourself of the last statement.)

PROBLEMS INVOLVING DIMENSIONS, UNITS AND VECTORS

Problems 1-11 are at the subject-test level. The remaining problems are at the AP level.

1. Copper has a density of nearly 9 grams per cubic centimeter. Expressed in terms of kilograms per cubic meter copper's density is:

   (A) $0.9\ \text{kg/m}^3$
   (B) $9\ \text{kg/m}^3$
   (C) $90\ \text{kg/m}^3$
   (D) $900\ \text{kg/m}^3$
   (E) $9000\ \text{kg/m}^3$

2. In the weekly lab, students were asked to weigh their physics texts to one percent. Which of the following would be the likely mass they measured?

   (A) 0.238 kg
   (B) 3.2 kg
   (C) 3.44 kg
   (D) 10.7 kg
   (E) 50 kg

3. Which of the following are vector quantities?

   a) Velocity b) Speed c) Displacement d) Mass e) Acceleration

   (A) a
   (B) a, b, e
   (C) a, d
   (D) a, c, e
   (E) b, d

4. If a vector **A** of length 5 is added to a vector **B** of length 4, the resultant vector will:

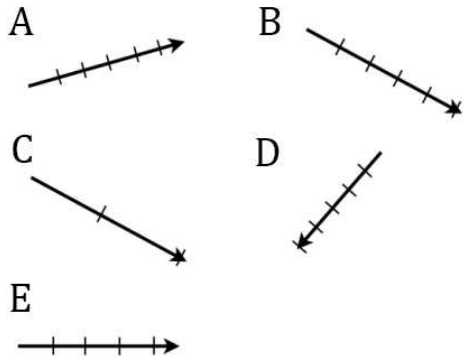
   (A) have a magnitude of 9.
   (B) have a magnitude of 1.
   (C) have a magnitude of -1.
   (D) have a magnitude somewhere between 0 and 9, inclusive.
   (E) have a magnitude somewhere between 1 and 9, inclusive.

5. A cyclist rides a bike 6 kilometers north and 9 kilometers east. Which is nearest to the final distance from her starting point?

   (A) 3 km
   (B) 9 km
   (C) 10 km
   (D) 16 km
   (E) 20 km

6. The same cyclist continues her trip by riding south for a distance of 11 km. Which of the below vectors best illustrates her final position from her starting point. Each tick mark represents two km; to the right is east and up is north.

   (A) A
   (B) B
   (C) C
   (D) D
   (E) E



7. Pressure is defined as a force per unit area. Which of the following has the same dimensions as pressure?

   (A) Energy per unit volume
   (B) Momentum per unit area
   (C) Momentum per unit length
   (D) Acceleration per unit area
   (E) Momentum per unit volume

8. A pendulum has a length ℓ, a mass m and is swinging under the action of Earth's gravity, which produces an acceleration g. Expressed in terms of these quantities, the pendulum's frequency, in cycles per second, could be:

   (A) $\sqrt{mg/\ell}$
   (B) $\sqrt{\ell/g}$
   (C) g/ℓ
   (D) $\sqrt{\ell g/m}$
   (E) $\sqrt{g/\ell}$

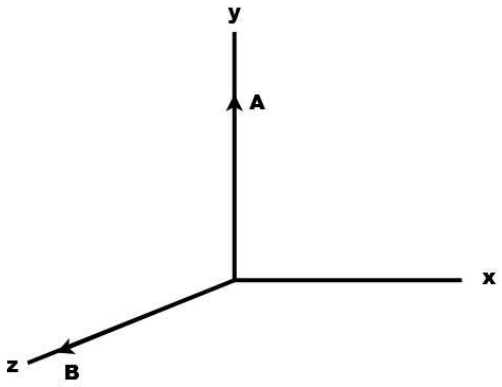
9. Power is defined as energy per unit time. Which of the following also represents power?

   (A) Momentum times acceleration
   (B) Momentum times velocity
   (C) Force times distance
   (D) Force times velocity
   (E) Both (A) and (D)

10. A sphere of mass M and initial radius R is compressed so that its final radius is $R/2$. How does the density change?

   (A) It stays the same.
   (B) It decreases by a factor of two.
   (C) It increases by a factor of two.
   (D) It increases by a factor of eight.
   (E) It increases by a factor of four.

Questions 11 - 12 refer to the following figure.



11. If **A** is a vector pointing along the y-axis and **B** is a vector pointing along the $+z$-axis, as shown above, then the vector $\mathbf{A} \times \mathbf{B}$

   (A) points along the $+y$-axis.
   (B) points along the $+x$-axis.
   (C) points along the $+z$-axis.
   (D) points along the $-x$-axis.
   (E) points along the $-z$-axis.

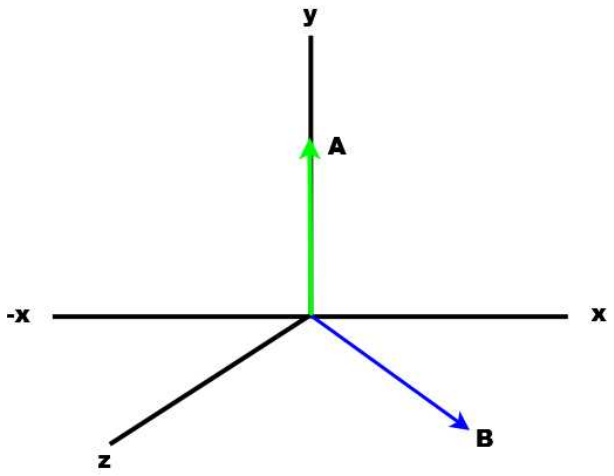
12. If **A** has length 2 and **B** has length 3, then $\mathbf{A} \times \mathbf{B}$ has length

   (A) 1
   (B) 2
   (C) 4
   (D) 6
   (E) 8

13. **A** is a vector with magnitude 3 pointing along the y-axis. The vector **B** has magnitude 4, lies in the xz-plane and points at 45° between the $+z$-axis and the $+x$-axis, as shown below.

The vector $\mathbf{B} \times \mathbf{A}$

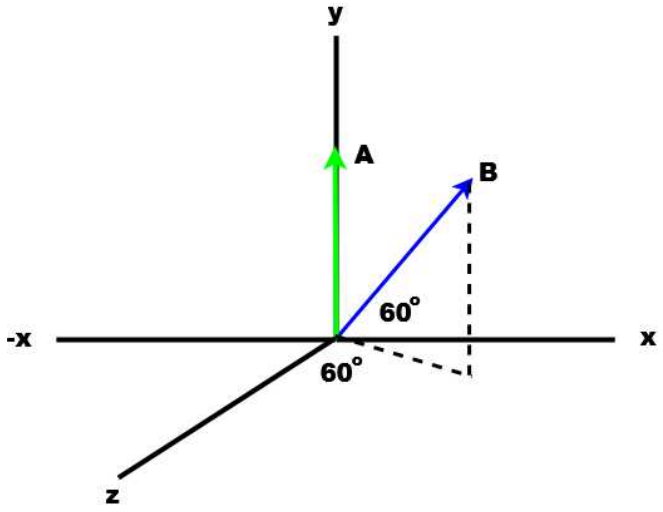
(A) has magnitude 12, lies in the xz-plane, and points at 45° from the $+z$-axis to the $-x$-axis.
(B) has magnitude 6, lies in the xz-plane, and points at 45° from the $+z$-axis to the $-x$-axis.
(C) has magnitude 6, lies in the xz-plane, and points at 45° from the $+x$-axis to the $-z$-axis.
(D) has magnitude 12, lies in the xz-plane, and points at 45° from the $+x$-axis to the $-z$-axis.
(E) has magnitude 12, lies in the xz-plane, and points at 45° from the $+z$-axis to the $+x$-axis.



14. **A** is a vector with magnitude 3 pointing along the y-axis. The vector **B** has magnitude 4, points upwards at 60° from the xz-plane at an angle of 60° from the $+z$-axis, as shown, below.

The vector $\mathbf{B} \times \mathbf{A}$

(A) has magnitude 6, lies in the xz-plane, pointing at 60° from the $+x$-axis to the $-z$-axis.
(B) has magnitude 12, lies in the xz-plane, pointing at 30° from the $+z$-axis to the $-x$-axis.
(C) has magnitude 6, lies in the xz-plane, pointing at 30° from the $+z$-axis to the $-x$-axis.
(D) has a magnitude 12, lies in the xz-plane, pointing at 60° from the $+x$-axis to the $-z$-axis.
(E) has a magnitude 6, lies in the xz-plane, pointing at 60° from the $+z$-axis to the $-x$-axis.



15. If **A** and **B** are vectors, then $\mathbf{B} \cdot (\mathbf{B} \times \mathbf{A})$ is equal to:

(A) **A**
(B) **B**
(C) A^2B
(D) zero
(E) B^2A

16. If **A** and **B** are vectors with an angle less than 90° between them, then we can say:

(A) $|\mathbf{A} + \mathbf{B}| = |\mathbf{A}| + |\mathbf{B}|$
(B) $|\mathbf{A} + \mathbf{B}| = \sqrt{|\mathbf{A}|^2 + |\mathbf{B}|^2}$
(C) $|\mathbf{A} + \mathbf{B}| < \sqrt{|\mathbf{A}|^2 + |\mathbf{B}|^2}$
(D) $|\mathbf{A} + \mathbf{B}| < |\mathbf{A}| + |\mathbf{B}|$
(E) (C) and (D)

17. If $\mathbf{A} = 3\,\hat{\mathbf{i}} + \hat{\mathbf{j}}$ and $\mathbf{B} = 6\hat{\mathbf{i}} - 3\hat{\mathbf{j}}$, then a unit vector in the same direction as $\mathbf{B} - \mathbf{A}$ is:

(A) $\hat{\mathbf{C}} = 9\hat{\mathbf{i}} - 4\hat{\mathbf{j}}$
(B) $\hat{\mathbf{C}} = -1/5(3\hat{\mathbf{i}} + 2\hat{\mathbf{j}})$
(C) $\hat{\mathbf{C}} = 1/5(3\hat{\mathbf{i}} - 4\hat{\mathbf{j}})$
(D) $\hat{\mathbf{C}} = 1/5(3\hat{\mathbf{i}} + 4\hat{\mathbf{j}})$
(E) $\hat{\mathbf{C}} = 1/84(9\hat{\mathbf{i}} - 2\hat{\mathbf{j}})$

18. If $\mathbf{A} = \frac{1}{5}(3\hat{\mathbf{i}} + 4\hat{\mathbf{j}})$ and $\mathbf{B} = \frac{1}{13}(5\hat{\mathbf{i}} + 12\hat{\mathbf{j}})$, then the angle between these two vectors is nearest

(A) 10.3°
(B) 14.3°
(C) 15.8°
(D) 25.7°
(E) 30.4°

2. KINEMATICS REVIEW

This entire chapter is at the subject-test level, although some problems are at the AP level.

1. Essential Terms

Mechanics is the branch of physics that deals with the motion of material objects under the action of **forces**. The term **kinematics**, which literally means "the study of motion," is the branch of mechanics in which the forces causing the motion are ignored; we merely consider the motion itself, not the causes.

The basic quantities under consideration in kinematics problems are **displacement**, **velocity** and **acceleration**. They are all **vector** quantities and must be treated as such.

2. Displacement and Velocity

Displacement is defined as the *position of an object from its point of origin after moving a time t.* Precisely, *displacement is the* ***vector*** *that points from the point of origin to the position of an object after a time t* (See Figure 2.1). The route the object took to get there is not important. People often speak of displacement and distance as being the same, but they are not. In making a round trip from New York to Philadelphia, you traverse a *distance* of about 300 kilometers, but your total *displacement* is zero. The dimensions of displacement and distance are length, L, and the SI unit of both is the meter, m. Displacement is often indicated by the letter **s**. Distance is a scalar and is usually indicated by Latin letters in italics, such as s or x.

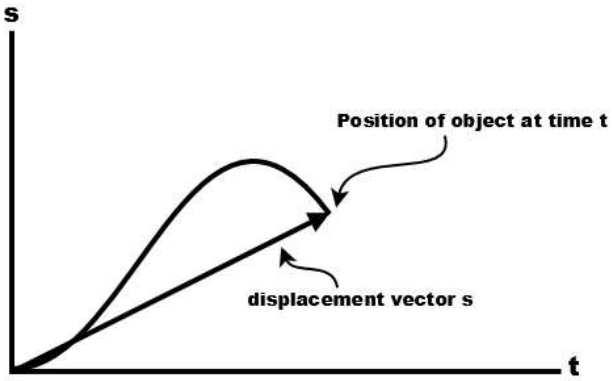


Figure 2.1

Velocity is defined as the *change in displacement per unit time*, or *the rate of change of the displacement vector*. **Speed** is the *magnitude of the velocity*, or the length of the velocity vector; speed is a scalar. The dimensions of velocity and speed are L/T. In SI units, velocity and speed are measured in meters per second, m/s.

Recall that the *slope* of a graph (rise over run) of any function represents the rate of change of that function. Also recall that the slope of a graph at any point is by definition the slope of the tangent line to the graph at that point.

Supreme Hint:

The slope of an object's displacement-versus-time graph at any point (the slope of the tangent at that point) gives the velocity of the object.

Commentary on hint: In general, the displacement may be a complicated function of time, as in Figure 2.2.

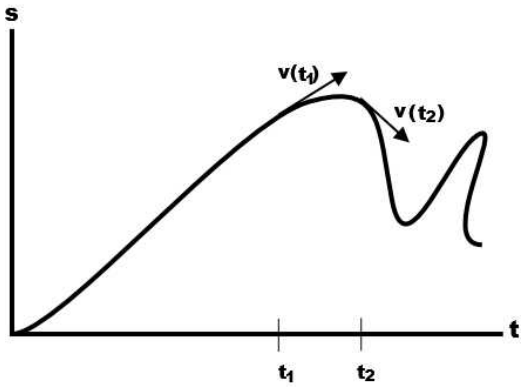


Figure 2.2

Because the slope of the graph changes continually with time t, the velocity must change at every point. To precisely find the velocity at each time requires calculus,[1] but over small time intervals we can approximate the velocity by drawing short tangent lines along the curve. Putting an arrowhead on a tangent line makes it a vector. Two such velocity vectors are shown in Figure 2.2. *The magnitude of the slope of the tangent vector gives the speed*. We calculate the slope, as always, by finding the rise over the run. The *rise* is $\mathbf{s}_f - \mathbf{s}_o$, the difference between the final and initial displacements (see figure 2.3, below). The *run* is $t_f - t_o$, the difference between the final and initial times. The velocity is

$$\mathbf{v} \equiv \frac{\mathbf{s}_f - \mathbf{s}_o}{t_f - t_o} = \frac{\Delta \mathbf{s}}{\Delta t} \quad (2.1)$$

The last expression above is read as "Delta $\mathbf{s}$ over Delta t", where the Greek letter Δ almost always stands for a small change in something. (**Exercise**: On Figure 2.2 draw two displacement vectors $\mathbf{s}_1$ and $\mathbf{s}_2$ that give one of the indicated velocity vectors.)

We have written the velocity as a vector, as one should. **In most College Board problems, motion is in one dimension. In that case, there are only two directions, positive and negative. We can then write v as a scalar and indicate $+v$ for velocity in the positive direction and $-v$ for velocity in the negative direction.**

As a one-dimensional example, suppose that an object's displacement $\mathbf{s}$ increases by 5 meters in the positive x-direction over a period of 2 seconds. We would say that the approximate velocity during the interval is $v = +(5/2)$ m/s. Precisely, this is the *average velocity*. We return to this point in Section 3.

Example 2.1: Finding approximate velocity from a displacement-vs.-time graph.

Take the displacement-vs.-time graph in Figure 2.2. We approximate the first hump (roughly) by straight-line segments over small intervals of time and draw Figure 2.3 below.

[1] Students who have taken calculus know that the instantaneous velocity, $\mathbf{v}(t) \equiv d\mathbf{s}/dt$, is the derivative of the displacement with respect to time.

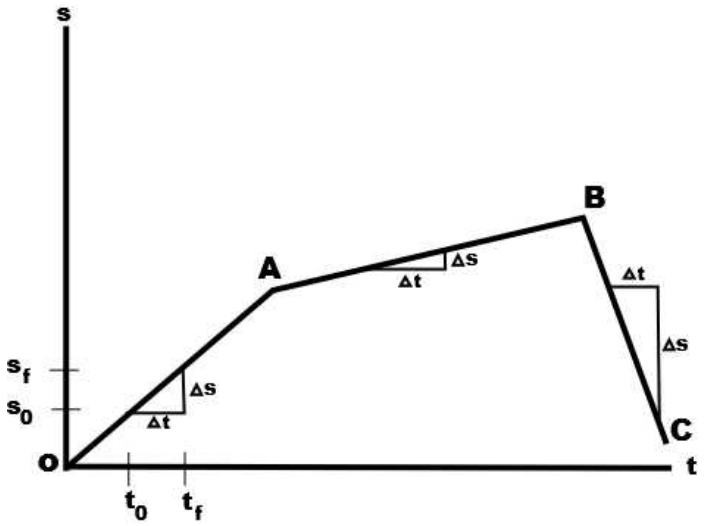


Figure 2.3

These lines segments approximate the tangents to the curve. In the interval **OA**, the slope $\Delta\mathbf{s}/\Delta t$ is positive, so the velocity **v** as defined by Eq. (2.1) is positive and the object is moving in the positive direction. In the interval **AB,** the slope $\Delta\mathbf{s}/\Delta t$ is smaller and positive, so the speed $v = |\mathbf{v}|$ is less but the object still has positive velocity and it continues moving in the positive direction. In the interval **BC**, the slope $\Delta\mathbf{s}/\Delta t$ is very large and negative, so $|\mathbf{v}|$ is very large but the object is moving in the opposite direction. The displacement, however, remains always positive. If point **A** has a displacement of 1 meter at a time of 2 seconds, then the velocity $\Delta\mathbf{s}/\Delta t$ in the interval **OA** is 0.5 m/s in the positive direction.

3. Relationship Between Total Displacement and Velocity

We have defined velocity as the *change* of displacement per unit time, $\mathbf{v} = \Delta\mathbf{s}/\Delta t$, which implies that a *small change* in displacement is $\Delta\mathbf{s} = \mathbf{v}\Delta t$. This statement applies only to small time intervals; it is a *common mistake* to say that the *total* displacement is $\mathbf{s} = \mathbf{v}t$ over a long period of time. This is strictly only true if the velocity is *constant*. Over an extended time interval, however, the velocity may be changing, just as when driving from New York to Philadelphia you may stop for gas or exceed the speed limit. Nevertheless, if someone asked you how fast you drove, you would say, "We drove 150 km in two hours, so we averaged 75 km per hour."

The colloquial statement that "velocity is displacement divided by time" is only approximately true. It is true, however, that the *average velocity*, $\bar{\mathbf{v}}$, is displacement/time, or $\mathbf{s} = \bar{\mathbf{v}}t$. Average velocity is usually written $\bar{\mathbf{v}}$ or $\langle\mathbf{v}\rangle$.

Here again, it is important to distinguish displacement from distance. If you made a round-trip journey from New York to Philadelphia in five hours, the average *velocity* would be zero, because you haven't gone anywhere—the total displacement is zero. On the other hand, the average *speed* would be $300/5 = 60$ km/hr. That is, the average speed is $\bar{v} = \overline{|\mathbf{v}|} = (\text{total distance})/\text{time}$.

Not coincidentally, there is an important graphical relationship between velocity and total displacement. Let us plot the velocities from Figure 2.3. Because the velocity is constant in each segment, one gets something like this:

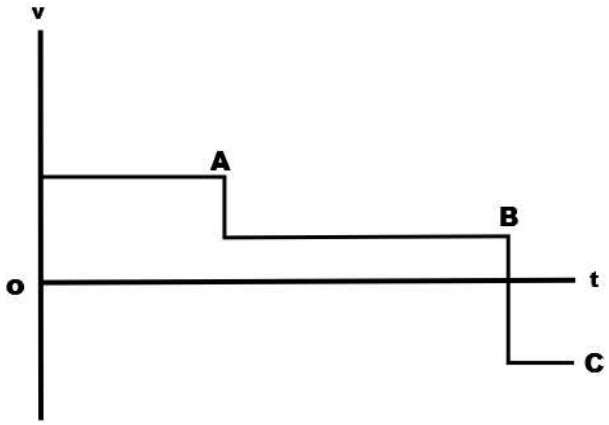


Figure 2.4

Notice that the total displacement, **s**, is the area on a velocity-vs.-time graph (the dimensions of displacement are $[\mathbf{s}] = \text{velocity} \times \text{time}$). If we calculated the area under the velocity curve in Figure 2.4, taking into account that the area in interval **BC** is counted as negative because it lies beneath the time axis, then the total area would equal the total displacement.

The area under a velocity vs. time graph is the total displacement.

Example 2.2: Finding total displacement from a velocity-vs.-time graph.

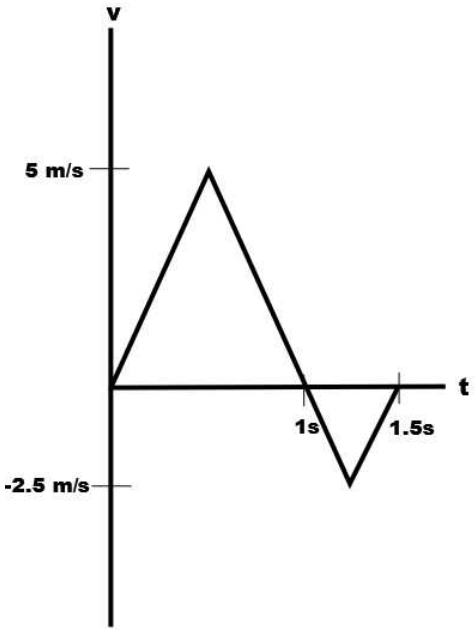


Figure 2.5

The area under the first triangle is $\frac{1}{2} \times 5 = 2.5$ m and the (signed) area under the second triangle is $\frac{1}{2} \times 0.5 \times (-2.5) = -0.625$ m. Adding them shows that the total displacement is $+1.875$ m.

The average velocity $\overline{\mathbf{v}}$ is in fact the velocity drawn on a velocity-vs.-time graph such that the area between it and the time axis equals the area beneath the actual velocity curve. In Figure 2.5, $\overline{v} = 1.25$ m/s because $(1.25 \text{ m/s}) \times (1.5 \text{ s}) = 1.875$ m.

Exercise: Verify the last statement and draw in $\overline{v}$ on Figure 2.5.

4. Acceleration

Acceleration is defined as the *change in velocity per unit time*, or *acceleration is the rate of change of velocity*. The dimensions of acceleration are $\mathrm{L/T^2}$. Since in the SI system, velocity is measured in $\mathrm{m/s}$, acceleration is measured by meters per second per second, $\mathrm{m/s^2}$.

As with velocity, the precise definition of acceleration at a given instant involves calculus.[2] However, in analogy to the definition of velocity, over small time intervals we can say

$$\mathbf{a} \equiv \frac{\Delta \mathbf{v}}{\Delta t} = \frac{(\mathbf{v}_f - \mathbf{v}_o)}{(t_f - t_o)} \quad (2.2)$$

where $\mathbf{v}_f$ is the final velocity, $\mathbf{v}_o$ is the initial velocity, and as before t_f is the final time and t_o is the initial time. If the velocity $\mathbf{v}$ changes from $\mathbf{v} = 5/2\ \mathrm{m/s}$ in the x-direction to $\mathbf{v} = 9/2$ meters per second in 2 seconds, the acceleration in that interval is $\mathbf{a} = \frac{9/2 - 5/2}{2}$ or $1\ \mathrm{m/s^2}$ in the x-direction. In analogy with the discussion in Section 3, this is technically the *average acceleration*. Also, in analogy with the discussion in Section 3, the area under an acceleration-vs.-time graph is the change in velocity.

Example 2.3: What is the acceleration in Figure 2.5?

The slope of the velocity is constant in each time interval in Figure 2.5, and so the acceleration is as well. The initial acceleration $\mathbf{a} = 5/0zx.5 = 10\ \mathrm{m/s^2}$ in the positive direction. The slope in the next interval, until the apex of the small triangle, has the same magnitude but is negative, so $\mathbf{a} = 10\ \mathrm{m/s^2}$ in the negative direction. The final acceleration is once again $\mathbf{a} = 10\ \mathrm{m/s^2}$ in the positive direction

5. Motion for Constant Acceleration

Just as the proper definition of acceleration in terms of velocity requires calculus, the reverse is also true. If an object's acceleration is changing with time, such that if $\mathbf{a} = \mathbf{a}(t)$, then it may be difficult, if not impossible, to find the exact velocity.[3]

Tip: Virtually all high-school and College Board problems assume that a is constant.

This assumption makes life much easier. For constant $\mathbf{a}$, the definition of acceleration in Eq. (2.2) equals the average acceleration for all time intervals and by cross-multiplying we can write

$$\mathbf{v}_f - \mathbf{v}_o = \mathbf{a}(t_f - t_o) \quad (2.3)$$

Usually we take the initial time to be zero, in which case t_f is just the total time t. Then

$$\mathbf{v}_f = \mathbf{a}t + \mathbf{v}_o \quad (2.4)$$

Thus, for constant acceleration, the velocity is just the acceleration multiplied by the time, as is commonly said, plus any initial velocity. Plotting Eq. (2.4) gives us the following picture.

---

[2] In analogy with velocity, $\mathbf{a}(t) \equiv d\mathbf{v}/dt$; acceleration is the derivative of the velocity with respect to time.
[3] In general, velocity is the integral of the acceleration with respect to time, $\mathbf{v}(t) = \int \mathbf{a}(t)dt$.

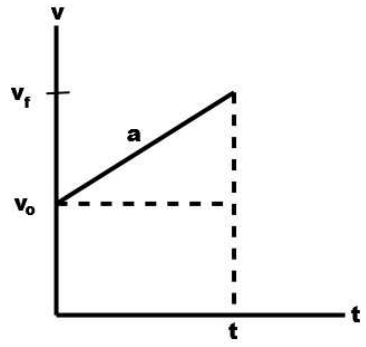


Figure 2.6

The area under this curve is $1/2(\mathbf{v}_f - \mathbf{v}_o)t + \mathbf{v}_o t = 1/2(\mathbf{v}_o + \mathbf{v}_f)t$. From Section 3 we know that this area equals the total displacement $\mathbf{s} = \bar{\mathbf{v}}t$, which implies that the object's average velocity is just $\bar{\mathbf{v}} = (\mathbf{v}_f + \mathbf{v}_o)/2$.

Using Eq. (2.4) for $\mathbf{v}_f$, we have

$$\mathbf{s} = \bar{\mathbf{v}}t = \left[\frac{\mathbf{v}_f + \mathbf{v}_o}{2}\right]t = \left[\frac{\mathbf{a}t + \mathbf{v}_o}{2}\right]t + \frac{\mathbf{v}_o}{2}t\,,$$

or

$$\mathbf{s} = \tfrac{1}{2}\mathbf{a}t^2 + \mathbf{v}_o t \quad (2.5a)$$

Note that the displacement $\mathbf{s}$ is the final position minus the initial position. Thus, if we let $\mathbf{s} = \mathbf{x} - \mathbf{x}_o$ where $\mathbf{x}$ and $\mathbf{x}_o$ represent the final and initial position vectors, respectively, Equation (2.5a) can be written

$$\mathbf{x} = \tfrac{1}{2}\mathbf{a}t^2 + \mathbf{v}_o t + \mathbf{x}_o \quad (2.5b)$$

This form is often encountered when the initial position is explicitly stated as not equal to zero.

Equations (2.4) and (2.5) are vector equations. If the motion is in only one dimension, we can write $\pm a$, $\pm v$ and $\pm s$ for motion in the positive (negative) direction, as was already discussed. Then, substituting $t = (v_f - v_o)/a$ from Eq. (2.4) into Eq. (2.5a) gives

$$s = \frac{1}{2}\frac{a(v_f - v_o)^2}{a^2} + v_o\frac{v_f - v_o}{a}$$

or, squaring the expression in parentheses and cancelling terms,

$$2as = v_f^2 - v_o^2 \quad (2.6)$$

Eq. (2.6) involves only scalar quantities (distance and speeds). Equations (2.4), (2.5) and (2.6) constitute the basic kinematic equations and are all you need to know to solve virtually any SAT subject question on kinematics. Memorize them**.**

Example 2.4: One–dimensional motion. A ball is thrown straight up into the air. What is the maximum height reached? What is the time required to reach the maximum height?

This is a simple application of Eqs. (2.4) and (2.5) and is used all the time. Once the ball leaves the hand, the only acceleration is that due to the Earth's gravity, $g = 9.8\ \text{m/s}^2$ acting in the negative (downward) direction. The maximum height y_m reached by the ball is attained when the acceleration due to gravity brings the velocity to zero. Thus, using Eq. (2.4), we get $0 = v_o - gt$. Hence, $t = v_o/g$. Plugging this result into Eq. (2.5) gives for the maximum height $y_m = -\frac{1}{2}g\left(\frac{v_o}{g}\right)^2 + v_o(\frac{v_o}{g})$, or $y_m = \frac{1}{2}v_o^2/g$. If you throw a ball up at $10\ \text{m/s}$ this means the time to reach the maximum height is $t = 10/9.8 \approx 1\ \text{s}$. The maximum height is $y_m = \frac{1}{2}100/9.8 \approx 5\ \text{m}$.

Exercise: If a ball is thrown upward, show that the time to reach the maximum height is the same as the time to fall from the maximum height to the ground. Thus, the roundtrip time is $t = 2v_o/g$. Show also that the speed with which the ball hits the ground is the same as the speed with which it is thrown upward.

Example 2.5: Two–dimensional motion. A projectile is fired at an angle $\theta = 60°$ from the horizontal at an initial speed $v_o = 50\ \text{m/s}$, as shown below. What is the horizontal range R attained by the projectile? Ignore any air resistance.

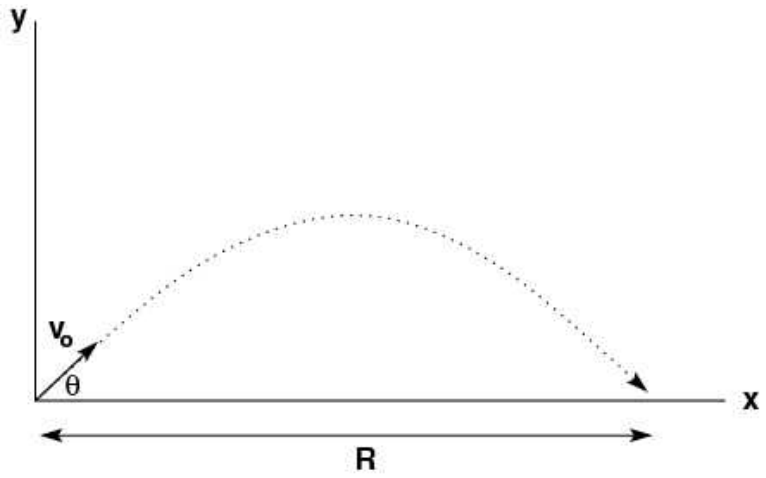


Figure 2.7

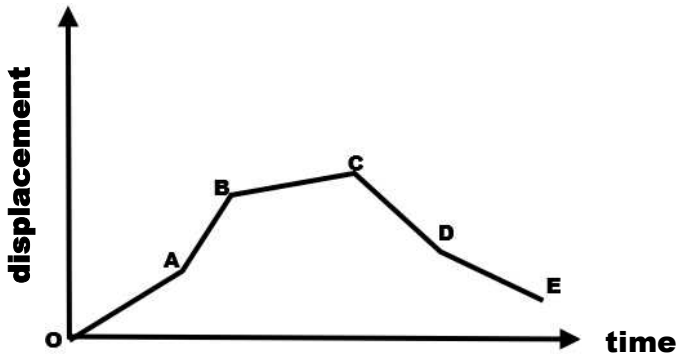
This is a basic problem whose result is used in any number of college exam questions. The important point is: *velocity is a vector*. Because the problem is two-dimensional, the velocity has both x- and y-components that can and must be treated separately. We write $v_o = v_{ox}\,\hat{\mathbf{i}} + v_{oy}\,\hat{\mathbf{j}}$, where v_{ox} is the initial speed in the x-direction and v_{oy} is the initial speed in the y-direction. As stressed in Section 1.4, **any time you are presented with a problem in two dimensions you should immediately resolve the vectors into components. The components act independently of one another. You have essentially turned one problem into two separate problems.**

In this example, there is no acceleration in the x-direction, and so the x-velocity remains constant: $v_x = v_{ox}$. The range is then simply $R = v_{ox}t$, where t is the time necessary for the projectile to hit the ground. From Example 2.4 (and the following exercise), $t = 2v_{oy}/g$ because only the velocity in the y-direction determines the projectile's maximum height. Thus, $R = 2v_{ox}v_{oy}/g$. From the figure we see that $v_{ox} = v_o\cos\theta$ and $v_{oy} = v_o\sin\theta$. Therefore, $R = 2v_o^2\sin\theta\cos\theta/g$. Using the double-angle trig identity, one can rewrite this as $R = v_o^2\sin 2\theta/g$. For the given numbers, $R = 2500 \times \frac{\sqrt{3}}{2}/9.8 \approx 220\ \text{m}$. Notice that R is maximized when $\sin 2\theta = 1$. Consequently, we have the important result that, barring air resistance, the maximum range of a projectile is attained by firing it at 45°.

KINEMATICS PROBLEMS

Problems 1-16 are at the subject-test level. The remaining problems are at the AP level.

Questions 1 - 6 refer to the following information.



A driver gets into his car at home and drives off along a straight road for a time until he stops for a rest. His initial direction is positive.

1. In which segment is his velocity most positive?

   (A) OA
   (B) AB
   (C) BC
   (D) CD
   (E) DE

2. In which segment is his speed the greatest?

   (A) OA
   (B) AB
   (C) BC
   (D) CD
   (E) DE

3. At which point has he traveled the furthest total distance from his starting point?

   (A) A
   (B) B
   (C) C
   (D) D
   (E) E

4. At which point is he farthest away from his starting point?

   (A) A
   (B) B
   (C) C
   (D) D
   (E) E

5. In which segment is his velocity the most negative?

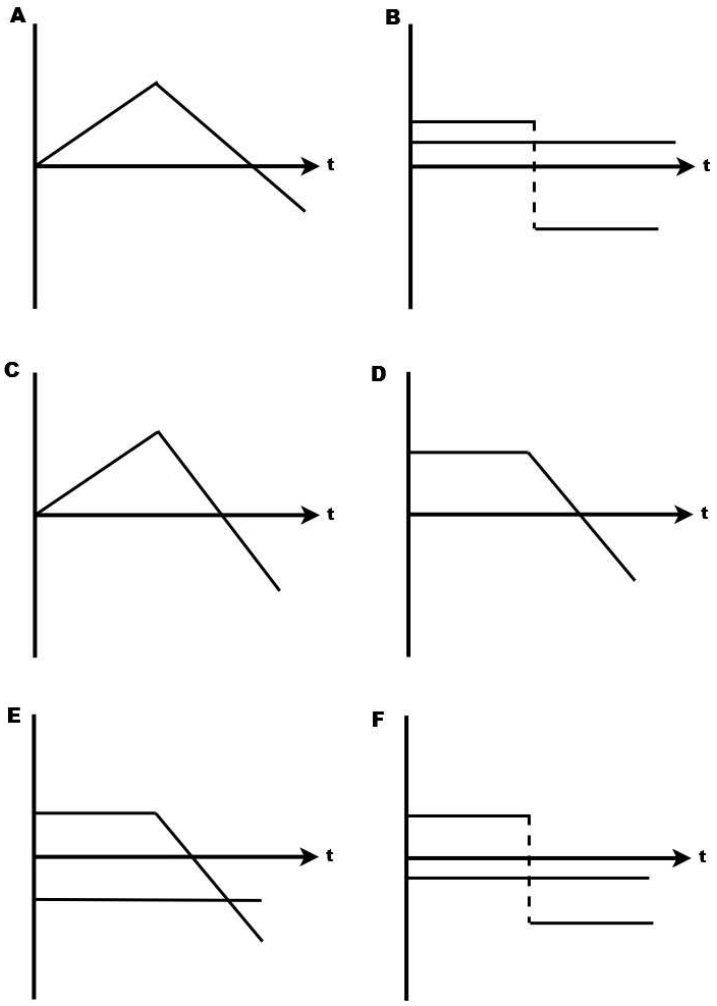
   (A) OA
   (B) AB
   (C) BC
   (D) CD
   (E) DE

6. In which segment is his acceleration the least?

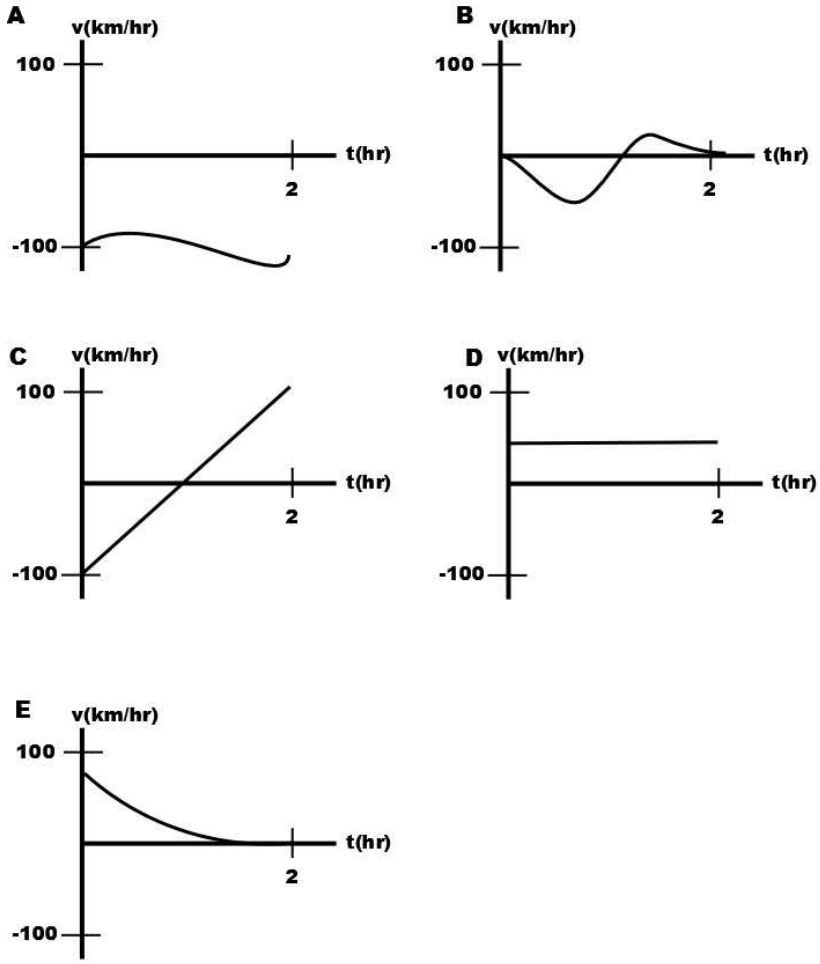
   (A) BC
   (B) DE
   (C) There is not enough information to solve the problem.
   (D) None
   (E) All the accelerations are equal.

7. A lady bug walks across a table in a straight line for 4 seconds, crossing 4 cm, then turns around and runs back in the opposite direction for 4 seconds, crossing 8 cm. Assume that the horizontal axes in all the below graphs represent time and that the scales are uniform. Which pair of graphs below could represent the lady bug's displacement and velocity, respectively, with her average velocity drawn in.

   (A) A and B
   (B) A and F
   (C) D and E
   (D) C and F
   (E) C and B



Questions 8 - 9 refer to the following figure.



8. Five drivers set off and drive for two hours on five straight roads. At the end of two hours, which car has the greatest displacement from its starting point?

   (A) A
   (B) B
   (C) C
   (D) D
   (E) E

9. Which two graphs have the closest average speeds?

   (A) A and D
   (B) A and C
   (C) C and D
   (D) C and E
   (E) B and D

10. A model rocket is launched into the air and after several seconds the engine has burned up all its propellant. Thereafter, the rocket continues to rise until it reaches a maximum height, a parachute opens and the rocket descends to the field. Assuming positive is up and negative is down, which are the correct signs of the displacement, velocity and acceleration during the ascent, after engine shutoff?

(A) A
(B) B
(C) C
(D) D
(E) E

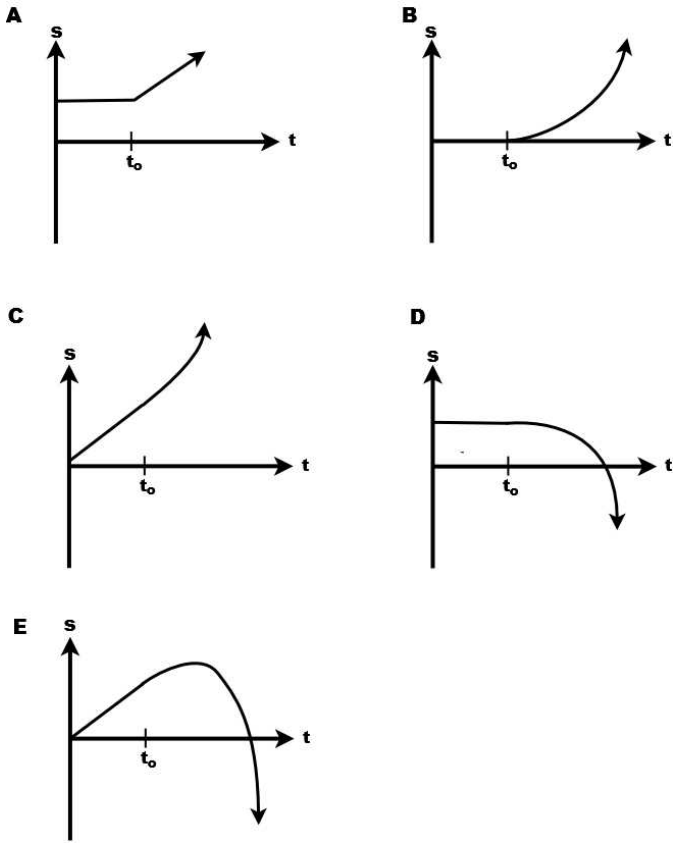
	Displacement	Velocity	Acceleration
A	Positive	Positive	Positive
B	Negative	Positive	Positive
C	Positive	Positive	Negative
D	Positive	Negative	Negative
E	Negative	Negative	Negative

11. Consider the same rocket as in the previous problem. What are the correct signs of the displacement, velocity and acceleration during the descent?

(A) A
(B) B
(C) C
(D) D
(E) E

12. A car sets off at time zero due north with constant initial velocity, $\mathbf{v}_o > 0$. At time t_o it begins to accelerate due south. Which of the below graphs could represent the car's displacement, with north taken as positive.

(A) A
(B) B
(C) C
(D) D
(E) E



13. A satellite in orbit around the Earth at a distance $R = 6800$ km from the Earth's center, completes 3 complete orbits in 4.5 hours. The satellite's average speed in km/s is:

(A) $\frac{2\pi(6800)}{1.5}\frac{1}{3600}$

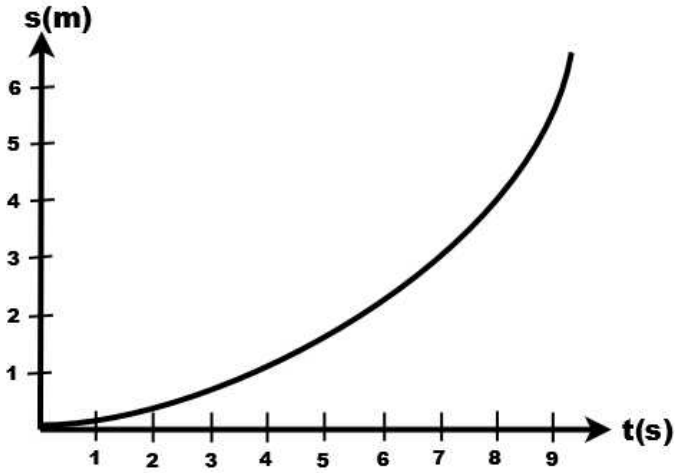
(B) $\frac{2\pi(6800)}{4.5}\frac{1}{3600}$

(C) $\frac{6800(3600)}{2\pi(4.5)}$

(D) $\frac{2\pi(3600)}{4.5(6800)}$

(E) $\frac{2\pi}{1.5(6800)}\frac{1}{3600}$

14. Shown below is a displacement-vs.-time graph for an object undergoing constant acceleration.

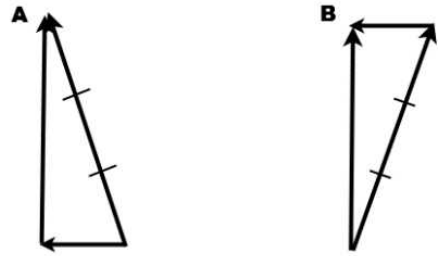


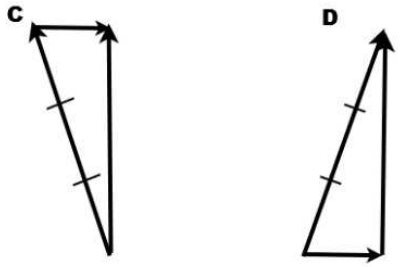
From this graph you can conclude:

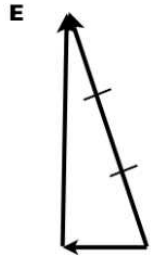
(A) that the average velocity between 4 and 6 seconds is approximately 0.5 m/s.
(B) that the average velocity between 7 and 8 seconds is about 1 m/s.
(C) that the acceleration is about $\frac{1}{8}\text{ms}^{-2}$.
(D) (A) and (C)
(E) (A), (B) and (C)

15. You wish to cross a river in a boat so that you end up on the opposite bank directly across from your starting point. You are able to row the boat at 3 km/hr in still water and the river is flowing to the left at 1 km/hr. The river is 1 km wide. Your strategy for crossing is best illustrated by which of the below diagrams?

(A) A
(B) B
(C) C
(D) D
(E) E





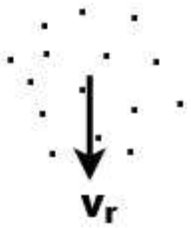


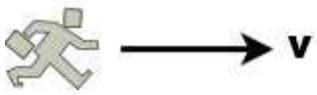
16. How long does it take you to cross the river in problem 15?

(A) $1/4$ hr
(B) $1/3$ hr
(C) $1/(2\sqrt{2})$ hr
(D) 1 hr
(E) $2\sqrt{2}$ hr

17. During a shower, rain is falling vertically at its terminal velocity v_r. A telecommuter, who is able to run very fast, is running horizontally to the right at a speed $v = \sqrt{3}v_r$. At what angle from the vertical does the telecommuter see the rain falling?

(A) 0° (straight down)
(B) 30°
(C) 45°
(D) 60°
(E) 90°

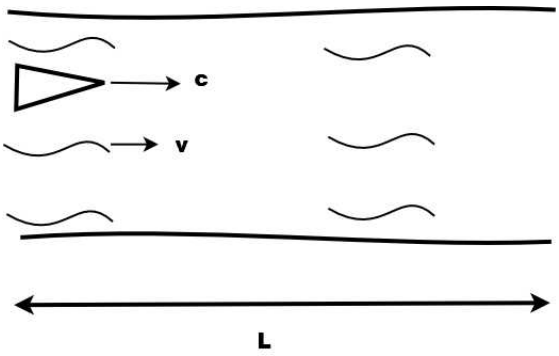




18. Suppose you try to head straight across the river of problem 15, rowing at 3 km/hr as you would in still water. How long does it take you to cross the river compared to the time it would take you in still water?

    (A) Cannot be determined from the information given
    (B) You cannot reach the other bank in a finite amount of time.
    (C) Less time
    (D) More time
    (E) The same amount of time

19. You can row a boat in still water at a speed of c km/hr. You decide to make a round trip in your boat down a river a distance L and back. The river is flowing to the right at a speed $v < c$, as in the figure below.



    The time it takes you to make the round trip is:

    (A) L/c
    (B) $2L/c$
    (C) Less than $2L/c$
    (D) Greater than $2L/c$
    (E) Cannot be determined from the information given

20. A ball is thrown off a building horizontally with a velocity of 10 m/s. The building is 45 m tall. Ignoring air resistance, the ball lands a horizontal distance from the building of approximately:

    (A) 90 m
    (B) 30 m
    (C) 20 m
    (D) 10 m
    (E) 3 m

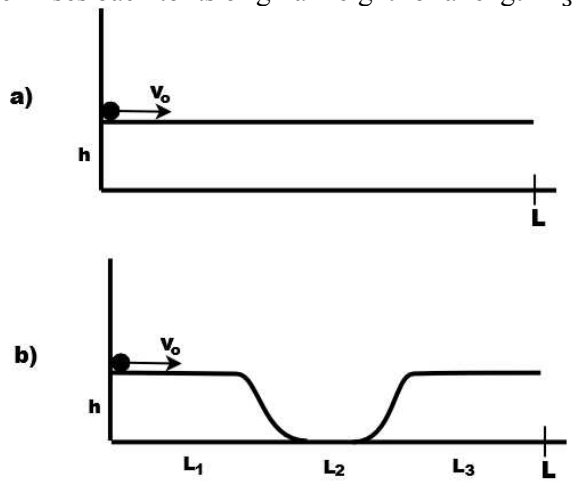
21. A raindrop has been falling downward through the air long enough so that it has reached a constant terminal velocity of 40 m/s. Over the next 3 seconds it will travel a further distance of:

    (A) 120 m
    (B) 180 m
    (C) 240 m
    (D) 320 m
    (E) 400 m

22. Two balls are pushed off a table of height h. Ball a has negligible horizontal velocity and falls straight down. Ball b has an initial horizontal velocity v_{ox}. Neglecting air resistance:

    (A) ball a hits the ground first.
    (B) ball b hits the ground first.
    (C) they both hit the ground at the same time, but a's speed is greater
    (D) they both hit the ground at the same time, but b's speed is greater.
    (E) they both hit the ground at the same time with the same speed.

23. As shown in the figure below, two balls are shot at initial velocity v_o along two tracks of horizontal length L and a height h above the floor. The first track is straight. After a length L_1, track b dips to the ground for a length L_2, then rises back to its original height for a length L_3.



    Ignoring any friction and the small curved sections of the track, which ball reaches the end of its track first?

    (A) Ball a
    (B) Ball b
    (C) They both reach the ends of their tracks at the same time with the same speed.
    (D) They both reach the ends of their tracks at the same time with different speeds.
    (E) The answer depends on the height h.

24. The conductor on a train spots a cow on the rails 200 meters ahead and immediately applies the brakes. The train decelerates at a constant deceleration a until it stops after 20 seconds, just before hitting the cow. Can the initial speed of the train in the instant before the brakes were applied be determined?

    (A) Yes, by dividing the given time by the given distance.
    (B) No, because to find the initial speed requires the deceleration, and the deceleration cannot be determined.
    (C) Yes, by solving for a in the equation $2as = v_o^2 - v_f^2$ and then using the equation $s = -\left(\frac{1}{2}\right)at^2 + v_o t$.
    (D) Yes, by finding the displacement in terms of the average speed.
    (E) Both (C) and (D).

3. Particle Mechanics Review

This is the most essential chapter of the book. As many as forty percent of SAT subject test questions concern mechanics and the concepts in all subsequent chapters rely on this material. Sections 3.7, 3.11, 3.13 and 3.14 are at the AP level. All the other sections are at the subject test level.

1. Essential Terms

As was mentioned in Section 2.1, **mechanics** is the branch of physics that describes the motion of material objects under the action of **forces**. In particular, **dynamics** refers to the branch of mechanics that explicitly deals with the forces causing accelerations, as opposed to **kinematics**, in which the forces causing the accelerations are taken for granted or ignored altogether. Dynamics is where the rubber meets the road and where as many as forty percent of SAT questions are concentrated.

Particle mechanics refers to the mechanics of "point objects" or objects that can at least be treated as point objects, such as billiard balls or blocks. The basic quantities under consideration in particle mechanics problems are **forces**, which affect **acceleration**, **velocity** and **displacement**. Equally important are **momentum** and **energy**. Above all are the **natural laws**, the laws governing how all these things—and more—behave.

2. Momentum

Momentum is one of the most fundamental quantities in physics and, loosely speaking, measures the "oomph" of a moving object or political campaign. Momentum is defined as the mass of the object multiplied by its velocity and for obscure reasons is usually represented by the letter **p**:

$$\mathbf{p} = m\mathbf{v} \quad (3.1)$$

Note that **momentum has a direction; it is a vector quantity and must be treated as such**. Since momentum is a mass times a velocity (momentum $= m\mathbf{v} \sim m\ell/t$), its dimensions are ML/T. No one's name has managed to get attached to momentum. In SI units it is measured in kg m/s.

Perhaps the most fundamental law of nature is:

Conservation of Momentum:

The total momentum of an isolated system never changes

In this case, "isolated" means not acted upon by an outside force. If one adds up all the products of all the masses and velocities of the bodies in such a system, their total momentum remains constant. The reason that conservation of momentum is so fundamental is that, in the first place, it is conserved even in situations when energy is not. Furthermore, like conservation of energy, it can be extended to systems that include non-material forms of momentum, such as electromagnetic waves. Of course, systems are often not isolated, but acted upon by outside forces, in which case we must consider

3. Newton's Laws and Forces

Governing how all material objects move are, of course, Newton's laws of motion:

I. All objects move at a constant velocity unless acted upon by an outside force.

II. Force is the rate of change of momentum.

III. For every action there is an equal and opposite reaction.

Commentary on the laws: The meaning of Newton's laws is not immediately apparent. Here is an interpretation:

1. Constant velocity includes zero velocity. Hence the popular alternative statement of the law, "An object at rest remains at rest, etc..." Basically, the first law says that all objects move in straight lines at constant speed unless acted upon by an external force. As a famous physicist once quipped, the first law states that any object moves at constant velocity insofar as it doesn't. Obviously the law doesn't make sense without the definition of a force. Newton I is also known as the **law of inertia**. Inertia is an imprecise term that means, loosely, "resistance to change in motion," but for objects of fixed mass it can be thought of as momentum and the law of inertia becomes equivalent to conservation of momentum. Newton's first law is due to Galileo.

2. Newton II is the definition of force. A force is any action that changes the momentum of an object. At its most basic, **a force is a push or a pull**. Common forces include gravity, the force of friction, the force exerted by a rope in pulling an object, the force exerted by a spring. Their behavior is defined by the second law. The true definition of a force is the change of momentum per unit time:

$$\mathbf{F} \equiv \frac{\Delta \mathbf{p}}{\Delta t} = \frac{\Delta(m\mathbf{v})}{\Delta t} \quad (3.2)$$

This equation forms the basis of all classical mechanics. **In virtually all high school and** College Board **problems, the mass of the object under consideration remains constant**, so Eq. (3.2) becomes $\mathbf{F} = m\Delta\mathbf{v}/\Delta t$, or the immortal

$$\mathbf{F} = m\mathbf{a} \quad (3.3)$$

Most students regard $\mathbf{F} = m\mathbf{a}$ as the definition of force; it is **NOT**. When the mass is constant, a force is anything that changes the velocity of an object—that is, causes an acceleration—but it should always be lodged in the back of one's mind that Eq. (3.3) is a special case of Eq. (3.2). If the mass is changing in the system, as in a rocket, one must use Eq. (3.2). It must also always be borne in mind that **F refers to the total external force on the object that is accelerating.** It does *not* refer to any force that the object may be exerting on another body. In solving mechanics problems, it is *crucial* to consider the force *on* the object under consideration. If all those forces cancel out, the object is said to be in **equilibrium** and is not accelerating. If the object is not moving at all, it is said to be in **static equilibrium**.

Pushes and pulls come with directions. Force is a vector quantity and must be treated as such. The dimensions of force are $[\mathrm{ma}] = \mathrm{ML/T^2}$. The SI unit of force is termed the newton and is defined as the force necessary to give a 1 kilogram mass an acceleration of 1 meter per second-squared: $1\mathrm{N} \equiv 1\mathrm{kg\ m/s^2}$.

3. Newton's third law is probably the most confusing of the three. The word "action" is an obsolete term for force. Newton III says that if object A exerts a force on object B, then object B exerts an equal and opposite force on object A. Thus, **forces always come in pairs**, known as **action-reaction pairs**. The confusion comes in deciding exactly what the pairs are. **The two forces of an action-reaction pair never act on the same body.** If I exert a force on you, you exert a force of equal magnitude but opposite direction on me. Period.

Example 3.1: Action-reaction pairs. Figure 3.1 indicates some action-reaction pairs for a person standing on a scale that is suspended by ropes from a ceiling above the Earth. By Newton III we must have $\mathbf{F}_{rc} = -\mathbf{F}_{cr}$, where $\mathbf{F}_{rc}$ is the downward force the rope exerts on the ceiling and $\mathbf{F}_{cr}$ is the upward force the ceiling exerts on the rope. **Exercise:** Identify the other action-reaction pairs.

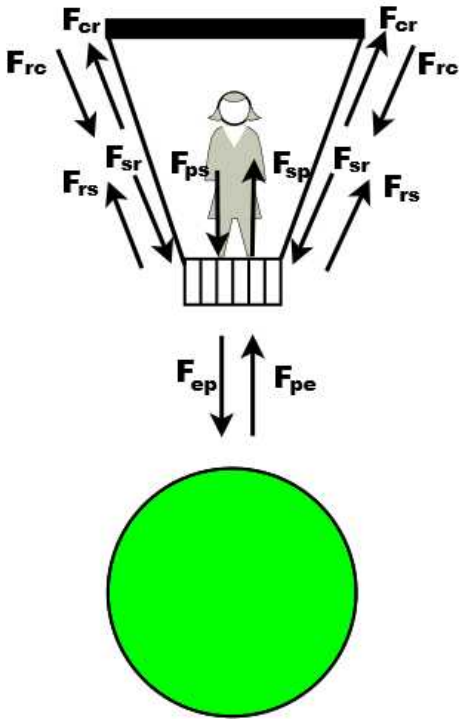


Figure 3.1

Example 3.2: Misapplying Newton's Laws. Consider two masses, M and m, in contact with one another and resting on a frictionless surface, as shown in Figure 3.2a, below. An external force, F**,** is applied to M from the left. In terms of F, what is the force that M exerts on m?

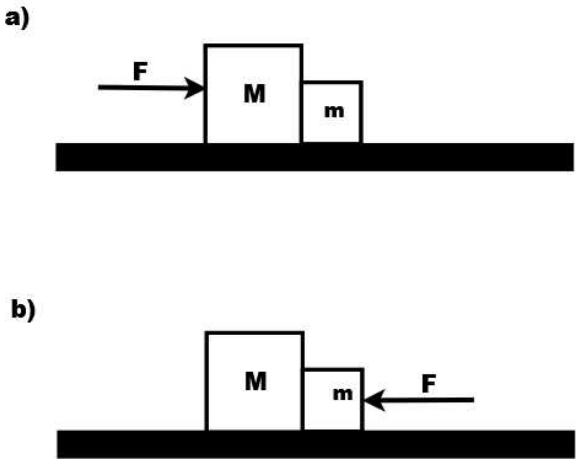


Figure 3.2

Many students say that F is "transmitted" by M to m and so the force on m is just F. Careful! The *only* force acting on m itself is exerted by M. Call this contact force F_{Mm}. Because m is the mass being accelerated, we must have $F_{Mm} = ma$. Now, since the two blocks are in contact they have the same acceleration, a. Newton II then requires that the external force F satisfies $F = (M + m)a$, implying $a = F/(M + m)$. Substituting this expression into the equation for F_{Mm} yields $F_{Mm} = mF/(m + M)$, which is *not equal* to F. The force F is *not* transmitted.

Furthermore, consider Figure 3.2b, where F is now applied to m from the right. The force exerted by m on M is $F_{mM} = -Ma$, implying that $F_{mM} = -MF/(M+m)$. Notice that $F_{mM} \neq -F_{Mm}$ from (a). Is this a violation of Newton III?

No! This is a completely new situation. F_{Mm} from (a) has nothing whatsoever to do with F_{mM} from (b). They are *not* action-reaction pairs. You can convince yourself that Newton III holds by considering (a) or (b) separately. The *total* force acting on M in (a) is

$$F + F_{mM} = Ma,$$

which by Newton III must be

$$F - F_{Mm} = Ma,$$

or substituting what we found above for a and F_{Mm},

$$F - \frac{mF}{m+M} = \frac{MF}{m+M}.$$

Cross multiplying shows $F = F$, which tells us that everything is correct. This example, although simple, shows how careful one must be when applying Newton's laws.

4. Solving Problems with Newton's Laws

In solving a mechanics problem, one always wants to know how an object does (or does not) move. The general prescription is:

0. Write down the total force *on* the object;
1. Find the acceleration from $\mathbf{a} = \mathbf{F}/m$;
2. From the acceleration find the velocity via Eq. (2.4);
3. From the velocity find the displacement with Eq. (2.5) or Eq. (2.6).

The zeroth task is generally the hardest. To write down the total force on the object one needs to identify *all* the forces acting on it. The prescription for finding the forces is: Draw a **free-body diagram**. The name may be mysterious but it is merely a picture of all the forces acting *on* the body under consideration. For example, if we wanted to know the forces acting on the person in Figure 3.1, there are only two: the force of the scale on her, $\mathbf{F}_{sp}$, which acts upward, and the force of the Earth's gravity, $\mathbf{F}_{ep}$, which acts downward. We might draw the free-body diagram below, which idealizes the person as a dot:

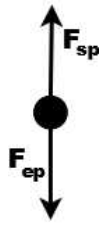


Figure 3.3

This diagram represents only the forces *on* the person; the other forces are irrelevant in terms of how *she* moves. Notice that the two forces *on her* are *not* action-reaction pairs. For her to remain motionless, the total force on her must be zero, so that $\mathbf{F}_{sp} = -\mathbf{F}_{ep}$; the gravitational force exerted on her by the Earth—her weight—must be equal and opposite the force the scale exerts on her. Any such force, which operates perpendicularly to a surface (in this case the scale) is termed a **normal force**. Your weight is equal to the normal force exerted by a scale on you.

Once you have accounted for all the forces, you can write down $\mathbf{a} = \mathbf{F}/m$. Before you can proceed further, however, you need to know the exact form of the forces produced by various phenomena.

5. Common Forces

Here is a list of forces commonly found in College Board problems and how they operate:

Gravity: For two spherical objects like planets, or for any objects far enough apart from each other, the force of gravity between them is given by

Newton's Law of Gravity

$$F_g = -\frac{Gm_1m_2}{r^2} \qquad (3.4)$$

In this equation, m_1 and m_2 are the two masses, r is the distance between their *centers* and G is a number, called the **gravitational constant**, which must be measured in the laboratory and sets the size of the force. In SI units, $G = 6.67 \times 10^{-11}\ \mathrm{m^3/(kg\,s^2)}$. Because gravity is a force, it must have a direction associated with it. *Gravity is always an attractive force, a pull, that acts along the line joining the centers of the two masses*. Newton III says that if m_1 is pulling m_2 to the right, then m_2 is pulling m_1 to the left with equal magnitude. (The $(-)$ in Eq. (3.4) indicates that gravity is an attractive force, but this is unimportant if one just wants to know the magnitude of the attraction, and one frequently writes the equation without the negative sign.)

The most important aspect of gravity is that it obeys an **inverse-square law**. *The force is inversely proportional to the square of the distance between the bodies*. This means that if you double the distance between the objects, the force decreases by a factor of four. Notice also that for estimates you can take $G = 10^{-10}\ \mathrm{m^3/(kg\,s^2)}$

Because Newton II remains operative ($F = ma$ to eternity!) we must have from Eq. (3.4) that $(Gm_1m_2)/r^2 = m_2a$, where a is the acceleration of m_2 produced by the gravitational attraction of m_1. It is a profound fact of nature that the m_2's cancel in this equation and **the gravitational acceleration $a = Gm_1/r^2$ does not depend on the mass of the body being accelerated, only on the mass causing the acceleration**. This is known as the principle of equivalence.

What is more, if $m_1 = M$ is the mass of the Earth and $r = R$ is its radius, then the gravitational acceleration at the Earth's surface is $GM/R^2 \equiv g$. As you certainly know, $g = 9.8\ \mathrm{m/s^2}$. The gravitational force the Earth exerts on any body at its surface is therefore

$$F_g = mg,$$

a quantity known as *weight*. Often we write $W = mg$. Weight, being a force, is technically measured in newtons. Colloquially, people speak of weight in pounds or kilograms. In physics problems you must remember to multiply such numbers by g to get the true weight. Also, in problems regarding Earth, gravity *always* acts straight down. **For almost any problem, it is permissible to take $g = 10\ \mathrm{m/s^2}$.**

Tension: When you pull on a car with a rope, you are exerting a force on the rope and the rope is in turn exerting a force on the car (Figure 3.4). That force is called a *tension*. With an ideal rope (one that is infinitely flexible) you cannot push on a car or anything else. **A tension is always a pull.** Moreover, in physics problems, "ideal rope" usually means one whose mass is negligible compared to all the other masses in the problem. *For a massless rope, the magnitude of the tension is constant throughout the rope* (see, eg., problem 3.20). In this case, the rope does transmit the force: if you pull on it with a force of 50 N, the rope pulls on the car with 50 N. Ropes can only support a maximum tension, above which they break. (**Exercise:** Identify the action-reaction pairs in Figure 3.4.)

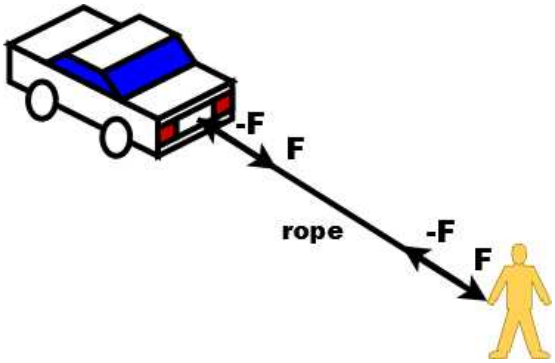


Figure 3.4

Friction: In the real world you know that friction operates by impeding the motion of an object. Precisely, **friction is a force that pushes oppositely to the direction of motion of a moving object**. In elementary physics one usually considers two types of friction: **kinematic friction** and **static friction**. Kinematic friction is the friction that impedes an object already in motion. Static friction is the friction that must be overcome to get an object moving. Technically, static friction pushes in the direction opposite that the object *would move*, once it got started. As a matter of experience, you know that static friction is greater than kinetic friction; once an object has begun moving, it is easier to keep it moving. Experimentally one finds that the magnitudes of both are proportional to the normal force N exerted by the surface upon which the object is resting or moving. The constant of proportionality is known as the **coefficient of friction** and is usually called μ ("mu"). Thus,

$$F_s \leq \mu_s N \text{ ; static friction} \quad (3.5a)$$

$$F_k = \mu_k N \text{ ; kinetic friction} \quad (3.5b)$$

(We have written only magnitudes here because the force of friction is not in the same direction as the normal force.) The $\leq$ appears in Eq. (3.5a) because $\mu_s N$ is the *maximum* force that static friction can exert; once this limit is attained, the object begins to move. In almost all elementary physics problems one is concerned with this maximum value. Often when we are not concerned whether we are talking about static or kinetic friction we write F_μ or simply f for frictional forces. Note that since both F_μ and N are forces, μ itself has no dimensions. It is a "pure number." Generally, μ runs between 0 and 1, depending on the materials involved, with $\mu_s > \mu_k$. Also, in most but not all elementary problems (see Example 3.3), the magnitude of the normal force is merely the weight of the object, and so one has $F = \mu Mg$ for an object of mass M.

Example 3.3: Friction and normal forces. In Figure 3.5a, a block of mass M rests on a horizontal surface. The normal force **N**, defined as the force exerted by the surface perpendicularly to itself, keeps the block from sinking into the surface and consequently must be equal and opposite to its weight, $M\boldsymbol{g}$. If the block is being pulled to the right by a force **F**, the frictional force **f** acts to the left. Its magnitude is $f = \mu Mg$, where μ is the coefficient of friction, and so the total force acting on the block is $F - \mu Mg$, to the right. On the other hand, if M is sliding down an incline, as in (b), the situation changes and one *must* resolve the vectors into their components (Section1.4).

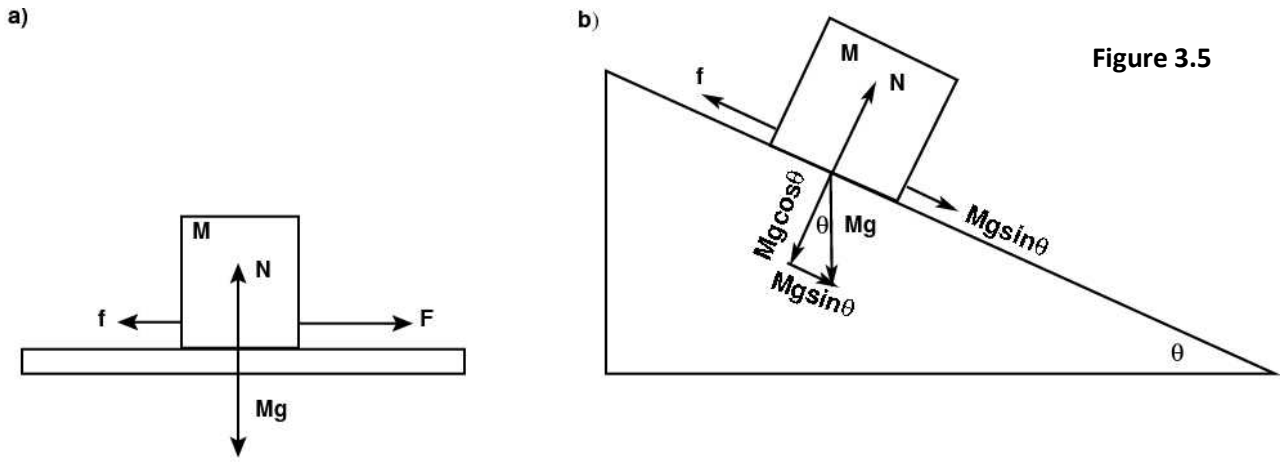


Figure 3.5

The force of gravity, Mg, always acts straight down. The component of gravity acting parallel to the incline is $Mg\sin\theta$, where θ is the angle at which the incline is tilted from the horizontal. The component perpendicular to the incline, $Mg\cos\theta$, must equal N in order that the block not sink into the surface. We see that in this case the normal force is *not equal* to the weight. Further, since $f = \mu N$, the total force in the parallel direction is $Mg\sin\theta - f = Mg\sin\theta - \mu Mg\cos\theta = Mg(\sin\theta - \mu\cos\theta)$. Notice that there is a critical angle θ_{crit} when $\mu = \tan\theta$. At θ_{crit} the total force on the block is zero and for angles less than or equal to θ_{crit}, friction prevents the block from sliding down the incline.

Spring force: Any spring exerts a force in the direction opposite to that which it has been either stretched or compressed. For small stretches or compressions, the magnitude of the force is experimentally observed to be directly proportional to the amount of stretch:

$$F = -k\Delta s \quad (3.6)$$

where Δs is the amount of stretch. Often this is written $F = -kx$, for a stretch in the x-direction, understanding that x represents the *change* in the length of the spring. Eq. (3.6) is known as **Hooke's Law**, although it is not a law of nature: if you stretch a spring past its elastic limit, it ceases to behave like a spring altogether. The constant of proportionality k is termed the **spring constant**. As you can see, k has dimensions force/length and is measured in newtons/meter.

We have written Eq. (3.6) as a scalar equation, but the $(-)$ indicates the direction: A spring force is termed a **restoring force**—it acts in the direction *opposite* to the force that streched or compressed it, tending to restore the system to its original state. Figure 3.6 illustrates this behavior: A mass is pulled to the left by a force F_p such that the spring is stretched by an amount Δs from its original length. The spring exerts a force $F = k\Delta s$ on the mass in the opposite direction, in this case to the right.

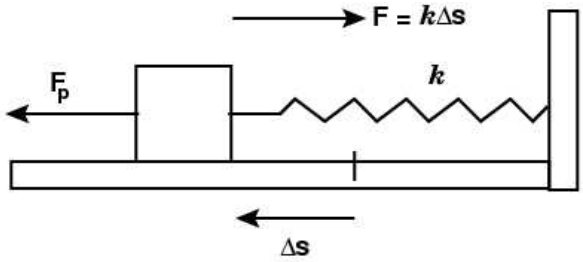


Figure 3.6

In elementary physics problems, an ideal spring is considered to be massless, and it is entirely characterized by the spring constant k. If it takes 100 N to stretch the spring 0.3 m, then $k = 333$ N/m. There is no typical value for a spring constant; they range from less than 1 N/m for a slinky to hundreds of N/m, depending on the stiffness of the spring. The stiffer the spring, the larger k.

6. Energy

Energy is one of the fundamental quantities of physics. As in life, in physics energy may be thought of as the ability to do work, more precisely to exert a force on an object across a distance. In mechanics we distinguish two types of energy: **kinetic energy**, *the energy of motion*, and **potential energy**, *the energy of position*. The sum of the kinetic and potential energy is referred to as **mechanical energy**. In thermodynamics we also speak of heat energy (which must be taken into account in systems with friction), although it is ultimately kinetic energy. In every situation, kinetic energy has the form

$$Kinetic\ Energy = KE \equiv \frac{1}{2}mv^2 \quad (3.7)$$

Note that kinetic energy depends on the *speed* of an object, which is why it is called the energy of motion. *Because only the square of the magnitude of the velocity appears in the definition, kinetic energy is a scalar* (yeah!). Its dimensions are ML^2/T^2 and in the SI system it is measured in kg m^2/s^2. **The dimensions of energy are the same as the dimensions of force × distance.** (**Exercise**: Convince yourself of this.)

The unit of energy in the SI system is known as the joule: 1 J ≡ 1 kg m^2/s^2. This is the same as 1 J = 1 N · m. Most contemporary texts use the symbol K for kinetic energy, but because there are too many k's floating around in physics we prefer KE for clarity.

As just mentioned, kinetic energy is always exactly as we have written it in Eq. (3.7). The complications invariably involve *potential energy*, the energy of position. This comes about because there is a close connection between potential energy and force. As there are many different kinds of forces—a few of which were listed in Section 5—there are consequently many forms for potential energy.[1] Potential energy is usually given the symbol U, but it has the same dimensions as kinetic energy (energy is energy) and is also measured in joules. For elementary courses and the College Board exams, the most important types of potential energy are *gravitational potential energy* and *spring potential energy*, also known to sophisticates as "harmonic oscillator potential energy."

$$Gravitational\ Potential\ Energy = U_g \equiv -\frac{Gm_1m_2}{r} \quad (3.8)$$

Here r is the distance between the *centers* of two spherical bodies. Notice that Eq. (3.8) differs from the expression for gravitational force in Eq. (3.4) only by the power of r in the denominator. That's a world of difference! If you remember that *the dimensions of energy are* $force \times distance$, one factor of r in the denominator of Eq. (3.4) must cancel when multiplying by *distance* and you will never get the two quantities confused.

Now, the acceleration of gravity on Earth is $g \equiv GM/R^2$, where M is the mass of the Earth and R is its radius. Multiplying g by the mass m of an object and a distance h must then give a potential energy. Indeed, for small distances above the surface of the Earth, one can write for the potential energy of a mass m:[2]

---

[1] Precisely, a force is the negative of the gradient of its potential, in one dimension, $F = -dU/dx$. The negative sign here explains why Eq. (3.4) needs to have a negative sign in front of it as well.

[2] Suppose a mass m rests a distance $h \ll R$ above the surface of the earth. From Eq. (3.8) its potential energy is $U_g = -\frac{GMm}{(R+h)} = -\frac{GMm}{R\left(1+\frac{h}{R}\right)}$. Multiply top and bottom by $(1-\frac{h}{R})$ to get $U_g = -\frac{GMm}{R\left(1-\frac{h^2}{R^2}\right)}(1-\frac{h}{R})$. If $h = 100$ m, then $h/R \approx 10^{-4}$ and $h^2/R^2 \approx 10^{-8}$, utterly negligible compared to 1. We can therefore say $U_g \approx -\frac{GMm}{R}\left(1-\frac{h}{R}\right) = -\frac{GMm}{R} + mgh$. The first term is the potential energy at the earth's surface, which may be taken as zero, and the second term is the additional energy due to the position of m at a height h above the surface.

$$U_g = mgh \text{ when } h \ll R \quad (3.9)$$

This is not the definition of gravitational potential energy, but it works well as long as the distance above the planet is much less than the radius of the planet itself (In Eq. (3.9) above, we are using the symbol $\ll$ to express this situation). This covers most daily situations and exam problems.

The fact that h in Eq. (3.9) represents a distance above a surface suggests that only differences in potential energy are important. Indeed,

Only energy differences can be measured. You can set the zero of energy where you want.

The other form of potential energy that frequently arises in elementary physics, is that associated with a spring. It is given by

$$U_s = \frac{1}{2}kx^2 \quad (3.10)$$

In this equation k is the spring constant and $x = \Delta s$ is the amount of stretch.[3]

7. Potential Energy Diagrams

It is imperative to have a mental image of the potentials discussed in Section 6. Figure 3.7 below illustrates them.

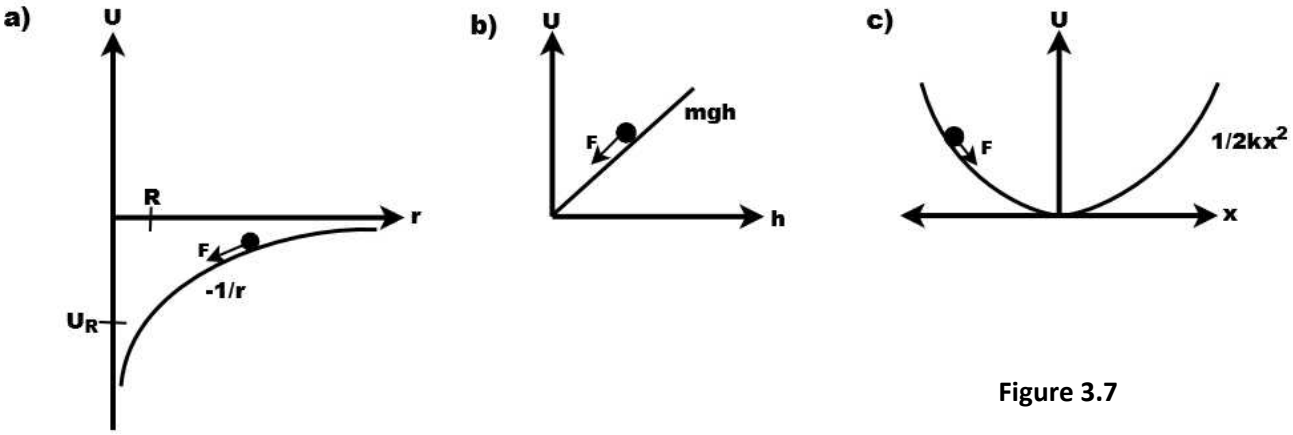


Figure 3.7

The most important thing about any potential is its shape. Sketch (a) shows the gravitational potential Eq. (3.8). It follows a $1/r$ curve and one immediately sees that $U \to 0$ as $r \to \infty$ and that $U \to -\infty$ as $r \to 0$. The radius of, for example, the Earth R is indicated as well as the potential U_R at that point, which is negative. Sketch (b) is the approximation to the potential (a) in the vicinity of R. U_R is taken to be zero at the point $h = R$ and the potential above that point is approximated as a straight line.

---

[3] Eq. (3.10) follows directly from the relationship between force and potential given in footnote 5. If the spring force is $F = -kx$, integrating with respect to x (and negating) gives Eq. (3.10).

Sketch (c) shows the spring, or harmonic oscillator, potential. Since Eq. (3.10) for the potential is $U = 1/2\, kx^2$, the shape is that of a parabola.

The direction of the force exerted on an object in the potential is easily seen from these diagrams. Merely place a ball somewhere on the potential; the direction it rolls is the direction of the force acting upon it. One can often learn many things just from looking at a potential energy diagram, without solving any equations. For instance, a glance at sketch (c) shows that the largest potential energy for a spring is for large x—when it is fully stretched—and the smallest potential energy occurs when $x = 0$, no stretch. Thus, by energy conservation (next section), $x = 0$ must be the point at which the kinetic energy is the largest. Also, you can see that Figure 3.8 below is not only a graph of height vs. position, but it is really a potential energy diagram—all you need to do is multiply the y-axis by mg. You conclude at once that in order to get over any of the humps shown, you must have an energy equal to mgh, where h is the height of the hump measured from wherever you are.

We cannot overstate how useful such diagrams are. To a physicist, a spring does not look like the picture in Figure 3.6, but the sketch in Figure 3.7c.

Example 3.4: Escape velocity. As a proof of the utility of potential energy diagrams, with the help of Figure 3.7a, it is easy to find the Earth's escape velocity—the velocity needed for a rocket to escape the Earth's gravitational pull. From Eq. (3.8) the rocket's potential at the surface of the Earth is $U_R = -GMm/R$, where m is the mass of the rocket. Figure 3.7a shows that as $r \to \infty$, the potential goes to zero. Thus, if we can give the rocket enough energy to raise its potential from U_R to zero, it will be a distance of infinity from the Earth—it will have escaped. We do this by giving the rocket a kinetic energy such that $\frac{1}{2}mv^2 + (-\frac{GMm}{R}) = 0$, which requires a velocity $v_{esc} = \sqrt{2GM/R}$. Plugging in the numbers for Earth, one finds that that v_{esc} is approximately $11\ \mathrm{km/s}$. This is the velocity you must give a rocket for it to completely escapes Earth's gravity.

8. Conservation of Energy

Most of the systems appearing in elementary physics courses, such as a mass falling to Earth or an oscillating spring, are termed **conservative systems**. As far as the College Board exams go, such systems are those without friction. Centuries of observation have led to the conclusion that **in conservative systems the sum of the kinetic and potential energy never changes**. In other words, **in conservative systems the mechanical energy is constant.** Friction generally slows objects down by changing kinetic energy into heat, but heat is another form of energy. Taking heat into account, the total energy in an isolated system—one disconnected from the rest of the world—always remains constant. This, of course, is a statement of the most famous law of nature,

Conservation of Energy:

The total energy of an isolated system never changes

For conservative systems, energy conservation requires that any change in mechanical energy is zero: $\Delta E = 0$. Therefore, we must have that any change in kinetic energy plus any change in potential energy equals zero:

$$\Delta \mathrm{E} = \Delta KE + \Delta U = 0,$$

or equivalently,

$$\Delta KE = -\Delta U \quad (3.11)$$

It is important to remember that Δ of anything *always* means (final – initial). Then, with $\Delta KE = \frac{1}{2}mv_f^2 - \frac{1}{2}mv_o^2$ and $\Delta U = U_f - U_o$, we can rewrite Eq. (3.11) as

$$\frac{1}{2}mv_f^2 + U_f = \frac{1}{2}mv_o^2 + U_o \quad (3.12)$$

If one holds a rock motionless in one's hand at a height h above the ground, $U_o = mgh$ and $KE_o = 0$. Letting the rock fall to the ground, $KE_f = \frac{1}{2}mv_f^2$ and $U_f = 0$. Consequently, $\frac{1}{2}mv_f^2 = mgh$. This is an extremely powerful statement. Solving Newton's equations can often be very difficult; to do so you need to know how the acceleration and velocity behave at every instant of time. But conservation of energy deals with only *the initial and final states of the system*, regardless of how the system got from one state to the other. This often makes seemingly impossible problems easy.

Example 3.5: Roller coaster. A ball rolling down a frictionless track, or a car on an ideal roller coaster, will speed up in certain places, and slow down in other places. To calculate the final velocity as the ball rolls from point A to point B would be difficult, conceivably impossible, using Newton's laws because the acceleration is constantly changing.

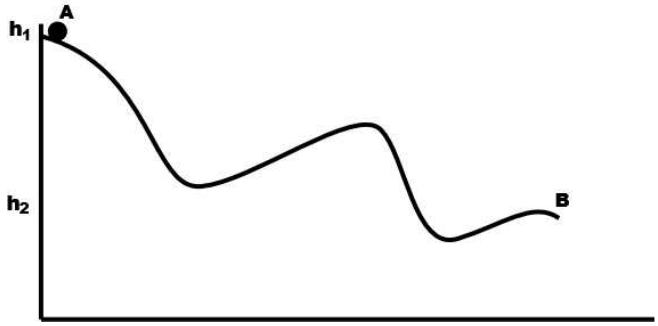


Figure 3.8

By conservation of energy, however, we can immediately say that $\frac{1}{2}mv_B^2 - \frac{1}{2}mv_A^2 = mgh_1 - mgh_2$, or that $v_B^2 - v_A^2 = 2g(h_1 - h_2)$. Notice that this is Eq. (2.6) with $a = g$.

Use conservation laws whenever possible

9. Work and Energy

That the final equation in Example 3.5 is the same as Eq. (2.6) is not coincidental. We have already pointed out that the dimensions of energy are the same as that of force × displacement. **The quantity force × displacement in physics is termed work**. The fact that the dimensions of work are the same as those for energy immediately suggests that there is an intimate connection between the two. To be sure, you know that lifting a large stone off the ground requires work, and that in the process you are also changing the stone's potential energy. If you lift the stone to a height h, then the force you are exerting up is equal and opposite the force that the Earth's gravitational field is exerting down. We therefore can say that the work you do against gravity is $W = F_g h = mgh$, where we recall that the force of gravity is $F_g = mg$. In one dimension, for a constant acceleration a and a displacement d,

$$W = mad \quad (3.13)$$

If work makes you mad, the formula for work is easy to remember.

Eq. (3.13) is the definition of work found in elementary texts. It is, unfortunately, an oversimplification and needs to be modified. First, while it may be easy to remember with a d, one must bear in mind that the relevant quantity is not distance but displacement, call it s. Second, the applied force only performs work if it actually displaces the object under consideration. A force acting straight down on the block in Figure 3.9 below does not move it anywhere and so it does zero work. Work is therefore properly defined in terms of the force parallel to the direction of displacement (see Example 3.6.). One can write

$$W = F_{\parallel} s \ .$$

This can be more conveniently written as

$$W = Fs\cos\theta = \mathbf{F}\cdot\mathbf{s} \quad (3.14)$$

where θ is the angle between the force and the displacement. (Those who have read Section 1.6 will recognize $\mathbf{F}\cdot\mathbf{s}$ as the dot product between the force and displacement. Those unfamiliar with the dot product can just use the expression involving cosine.)

Hint: If the force is perpendicular to the displacement, no work is performed.

Also note that if **F** is parallel to **s**, the work is positive, but if **F** is antiparallel to **s** the work is negative. Textbook writers unfortunately give the impression that positive and negative work has an absolute meaning, but the sign of the work is a convention; it depends on who is doing work on what. A more physical way to look at the situation is that if the work is positive, you are *putting energy into* the system; if the work is negative, you are *taking energy out* of the system.

Example 3.6: Work on a block. If a force **F** is applied to a mass m at an angle θ, as in Figure 3.9, only the horizontal component of the force displaces the block. Therefore, the work done through a displacement s is $W = Fs\cos\theta$. Note that if $\theta = 0$, $W = Fs = mas$, as in Eq. (3.13), but if $\theta = 90°$, $W = 0$.

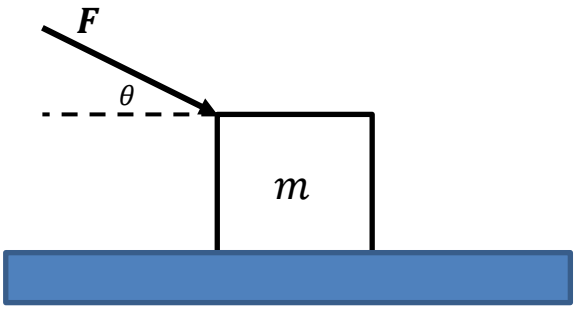


Figure 3.9

Even the statement $W = Fs\cos\theta$ is correct only for constant forces, which is generally the type encountered on the College Board exams.[4] Let us assume, then, a constant acceleration a in one dimension and that the acceleration is in the direction of the displacement s. With the help of Eq. (2.4) we can write

$$W = m\left(\frac{v_f - v_o}{t}\right)s.$$

We also showed in Section 2.5 that $s = \overline{v}t$, where $\overline{v} = 1/2\ (v_f + v_o)$. Thus,

$$W = \frac{1}{2}m(v_f - v_o)(v_f + v_o),$$

or

$$W = \frac{1}{2}mv_f^2 - \frac{1}{2}mv_o^2 \quad (3.15)$$

[4] It is more precise to say that an increment of work $\Delta W = \mathbf{F}\cdot\mathbf{\Delta s}$. (For those who know calculus, the correct expression is $dW = \mathbf{F}\cdot\mathbf{ds}$. To get the total work done for a nonconstant force, one must integrate this expression.)

which we recognize as $W = KE_f - KE_o \equiv \Delta KE$. **Work is the change in kinetic energy of an object that has undergone a displacement *s* produced by a given force F**. The statement is more general than this simple derivation suggests; it holds for nonconstant acceleration and motion in three dimensions as well.[5] The famous result should be called the **work-kinetic energy theorem**, but is almost always and misleadingly called just the **work-energy theorem**.

Work-Energy Theorem:

Work is the change in kinetic energy

For conservative systems, the work-energy theorem is essentially a statement of energy conservation. Eq. (3.11) tells us that *in the absence of friction* $\Delta KE = -\Delta U$. Thus, the work-energy theorem requires

$$W = -\Delta U \quad \text{(conservative systems)}.$$

That immediately implies that $W = \frac{1}{2}mv_f^2 - \frac{1}{2}mv_o^2 = U_o - U_f$, or $\frac{1}{2}mv_f^2 + U_f = \frac{1}{2}mv_o^2 + U_o$, energy conservation again. We also now understand that elementary definition of work Eq. (3.13), which equates work with a change in potential energy, can be true only for conservative systems. And finally, we now recognize that Eq. (2.6) is just an expression of the work-energy theorem for conservative systems.

Warning: The entire story changes for systems with friction—nonconservative systems (Section 11).

10. Power

In many situations, we would like to know the rate at which energy is being expended. To raise a refrigerator from one floor to another always requires the same amount of energy, mgh, but a larger motor is required if you have to do it in a short amount of time compared to a long amount of time. **Power** is defined as the rate of change of work:

$$P \equiv \frac{\Delta W}{\Delta t}$$

For a constant acceleration, we saw in Eq. (3.14) that $W = F_{\|}s = Fs\cos\theta$. Thus, $\Delta W/\Delta t = F\Delta s/\Delta t\ \cos\theta$. But $\Delta s/\Delta t \equiv v$, and so in analogy with Eq. (3.14), power becomes

$$P = F_{\|}v = Fv\cos\theta = \mathbf{F}\cdot\mathbf{v} \qquad (3.16)$$

where, as in Figure 3.9, θ is the angle between the applied force and the direction of motion. (For readers of Section 1.6, the dot product has again made an appearance. As before, other students can ignore it.) Note that **if the force is perpendicular to the direction of motion, no power is expended.**

The same remarks concern positive and negative power as previously concerned work. If **F** is parallel to **v**, you are putting energy into the system ("positive work"); if **F** is antiparallel to **v**, the force is acting oppositely to the direction of motion and you are taking energy out of the system ("negative work").

---

[5] The general result requires calculus: $W \equiv \int_o^f \mathbf{F}\cdot\mathbf{ds} = \int_o^f m\mathbf{a}\cdot\mathbf{ds} = m\int_o^f \frac{\mathbf{dv}}{dt}\cdot\mathbf{ds} = m\int_o^f \frac{\mathbf{dv}}{\mathbf{ds}}\frac{\mathbf{ds}}{dt}\cdot\mathbf{ds}$. But $\mathbf{ds}/dt \equiv \mathbf{v}$, so $W = m\int_o^f \mathbf{v}\cdot\mathbf{dv} = \frac{1}{2}mv_f^2 - \frac{1}{2}mv_o^2$.

Like energy, power is a scalar. Dimensionally, power = energy/time or force × velocity. The SI unit of power is the watt. 1 W ≡ 1 J/s or 1 N m/s.

Example 3.7: Suppose you want to install an elevator in the Freedom Tower that is 1000 kg fully loaded and can carry passengers to a height of 300 m in thirty seconds at constant velocity. What size motor do you need to purchase?

Assuming that the system is conservative, we must raise the elevator to a potential energy of $mgh = 1000 \times 10 \times 300 = 3 \times 10^6$ J. If we are to do this in 30 seconds, then $P = \frac{mgh}{t} = 10^5$ W. But according to Eq. (3.16), we should also have $P = Fv$. Note that because the elevator is being raised at a constant velocity, the tension T in the cable must equal mg, the force of gravity. Then $P = Tv = mgv = mgh/t$, the same as previously, 10^5 W. That might seem like a large number but in terms of a perhaps more familiar unit, it amounts to a 134 horsepower motor (1 hp = 745.7 W).

The careful student might well be puzzled by this example. We defined power as the change in work per unit time. Work, in turn, was defined as the change in kinetic energy. Here, because the elevator is moving at constant velocity, the total force on the elevator must be zero, and there is no change in kinetic energy. Why isn't the work zero? This is actually a subtle point. As far as the *motor* is concerned, it must exert a force of $T = mg$ on the elevator, and so the power it expends is indeed $P = Tv$, as just computed, which results in a total work of $W = Pt = Tvt = mgh$. The gravitational field of the Earth is performing the same amount of work in the opposite direction, so that the total is zero. In computing work, you must specify exactly what force you are talking about.

11. Work and Power in Nonconservative Systems

Suppose you are pushing a block up an incline such that there is friction between the incline and the block. *Since both you and friction are exerting a force through a displacement*, *both you and friction do work*. If W_c is the (conservative) work you would normally do in the absence of friction and W_μ is the work done by friction itself, then the work-energy theorem requires that $\Delta KE = W_c + W_\mu$. With $W_c = -\Delta U$ from above, $W_\mu = \Delta KE + \Delta U$. In other words,

$$W_\mu = \Delta E \qquad (3.17)$$

The work done by friction is equal to the change in the total mechanical energy of the system.

Example 3.8: Friction and mechanical energy.

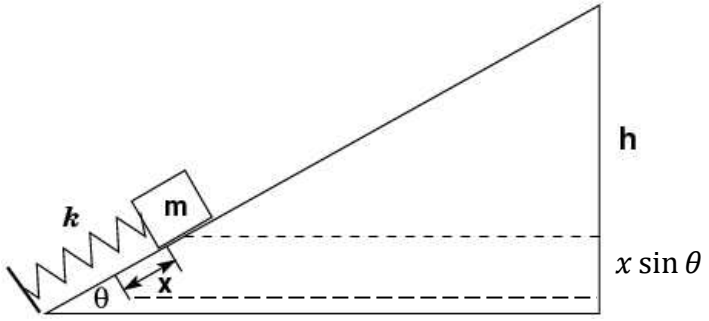


Figure 3.10

A spring with spring constant $k = 40$ N/m in contact with a mass $m = 25$ gm is compressed a distance of $x = 10$ cm *along* an incline that is tilted at an angle of 30° from the horizontal. The spring is released. Assume that m leaves the spring after traveling a distance x. From that point on, the incline is rough and the mass travels to a height of $h = 0.6$ m above the release point. What is the value of the coefficient of friction μ? You may take $g = 10$ m/s^2.

First, at the point at which the spring is fully compressed, the total energy available to m is just the spring's potential energy. From Eq. (3.10), this is $U = \frac{1}{2}\,kx^2$. The spring is released and the mass travels a distance x when it leaves the spring. At that point, the spring energy has been used to give the block a kinetic energy $1/2\,mv^2$, but has also been used to raise the block a height $x\sin\theta$ (see Figure 3.10), giving it a gravitational potential energy of $mgx\sin\theta$. By energy conservation we must have $1/2\,kx^2 = 1/2\,mv^2 + mgx\sin\theta$ or $\frac{1}{2}mv^2 = \frac{1}{2}kx^2 - mgx\sin\theta$. One can now solve for v, the velocity of the block when it leaves the spring, but this is unnecessary.

By conservation of energy, the height h_o that m should reach in the absence of friction is given by setting the change in potential energy equal to the change in kinetic energy: $mgh_o = \frac{1}{2}mv^2$, or substituting from above, $h_o = \frac{1/2\,kx^2}{mg} - x\sin\theta$. Plugging in all the numbers gives $h_o =$ 0.75 m. However, we are told that *with* friction the block reaches a height $h_\mu = 0.6$ m. The difference in energy $\Delta E = mg(h_o - h_\mu) \equiv mg\Delta h = 0.15/4$ J. This energy must have been dissipated by nonconservative forces—friction.

The work due to friction is $W_\mu = \mu N d$, where N is the normal force and d is the total displacement of m. From Example 3.3, we know that $N = mg\cos\theta$. Since d is the displacement *along* the slope, $d = h/\sin\theta$. Equating W_μ to ΔE gives $mg\Delta h = \mu mgh\cos\theta/\sin\theta$, or $\mu = (\Delta h/h)\tan\theta$. For the given numbers, we see that $\mu = (0.15/(0.6\sqrt{3}) = 0.144$, which is a plausible number for a coefficient of friction.

Exercise: Do the problem beginning from Eq. (2.6) and with the results from Example 3.3 show that it leads to the same answer.

Example 3.9: What is the average power expenditure by friction in Example 3.8?

We already know that friction dissipated $1/20$ J of energy. The only remaining question is how long it took to do this. The relevant time is the time it took for the mass to reach $h_\mu = 0.6$ m. The easiest way to find t is to use Eq. (2.5), $v_f - v_o = -at$, where we have inserted a $(-)$ to account for the fact that the acceleration is negative. Since the mass stops at its maximum height, $v_f = 0$; we have merely $v_o = at$, or from Example 3.3,

$$v_o = g(\sin\theta - \mu\cos\theta)t.$$

Plugging in the numbers already found in Example 3.8 and solving for t gives $t \approx 2.66$ s. Therefore, the average power expended by friction is $P = \frac{1}{20} \times \frac{1}{2.66} \approx 1/53$ W.

12. Collisions

In mechanics, probably the most important interactions between objects are collisions. Indeed, the entire subject of modern particle physics is based on analyzing collisions between subatomic particles. The most important feature of collisions is the following.

During any collision momentum is conserved.

The total momentum of the system after the collision equals the total momentum before the collision:

$$\mathbf{p}_f = \mathbf{p}_o$$

Moreover, because momentum is a *vector* and can be resolved into x- and y-components that behave independently, during a collision momentum in the x-direction is independently conserved and momentum in the y-direction is independently conserved. That is, $\mathbf{p}_{fx} = \mathbf{p}_{ox}$ and $\mathbf{p}_{fy} = \mathbf{p}_{oy}$. In three dimensions, the same holds for the z-component.

In basic physics we distinguish between two types of collisions: **elastic collisions** and **inelastic collisions**. An elastic collision, loosely speaking, is one in which the two objects "bounce off" of each other. *A perfectly elastic collision is defined to be one in which the kinetic energy before the collision is equal to the kinetic energy after*:

$$KE_f = KE_o \text{ (elastic collision)}$$

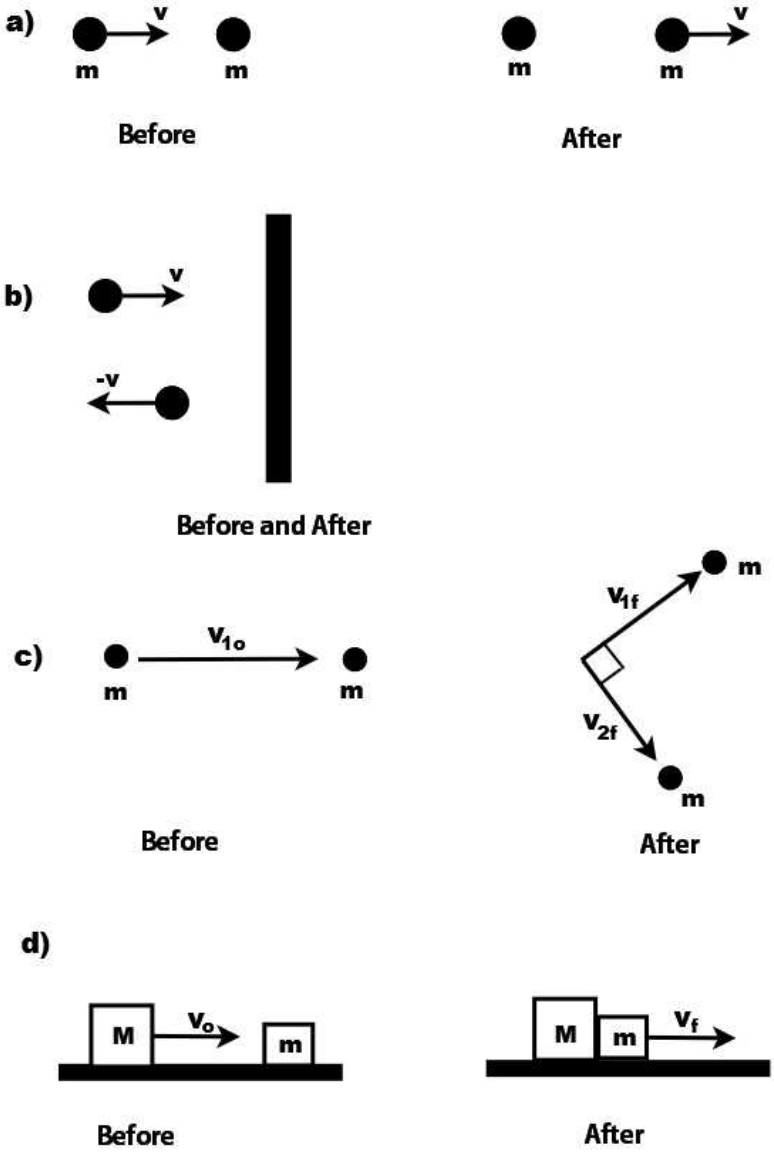
There is no law of conservation of kinetic energy. In elastic collisions kinetic energy is conserved by definition.

An inelastic collision, on the other hand, is one in which the colliding objects "stick together."

In inelastic collisions kinetic energy is not conserved.

Some of the energy is lost as heat, but **momentum is always conserved.**

Example 3.10: **Basic collisions.** The below figures illustrate the basic elastic and inelastic collisions. Except for (c) below, they are all one-dimensional. .



a) A mass m moving to the right at velocity v_{1o} collides elastically with an equal mass initially at rest. By conservation of momentum, $mv_{1o} = mv_{1f} + mv_{2f}$. Canceling the m's gives

$$v_{1o} = v_{1f} + v_{2f}.$$

Because this is an elastic collision, $KE_f = KE_o$, or

$$v_{1o}^2 = v_{1f}^2 + v_{2f}^2.$$

The first equation can be rewritten as $v_{1f} = v_{1o} - v_{2f}$. Inserting this expression for v_{1f} into the second equation, squaring everything and canceling yields

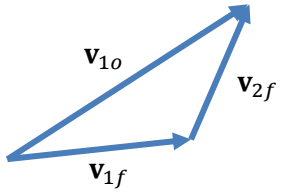
$$v_{2f}(v_{2f} - v_{1o}) = 0.$$

The only two possibilities are that $v_{2f} = 0$ or that $v_{2f} = v_{1o}$. The former doesn't make physical sense, because according to the top equation, that would mean $v_{1f} = v_{1o}$, or that no collision took place. Thus, we must have $v_{2f} = v_{1o}$ and therefore, $v_{1f} = 0$. *When a mass collides elastically with an equal mass at rest, the two masses exchange velocities*. You have frequently seen this result in action on the pool table.

b) A superball with initial momentum $p = mv$ collides elastically with a wall and bounces off in the opposite direction. Because kinetic energy is conserved, its final speed is equal to its initial speed, but the velocity has changed from $+v$ to $-v$. The ball's final momentum is thus $p = -mv$ and its change of momentum is $\Delta p = -mv - mv = -2mv$. Does this mean that momentum is not conserved? No. The wall, which is attached to the Earth, has gained a momentum of $2mv$, but the wall-earth system is so enormously more massive than the ball, that its recoil velocity is for all intents and purposes zero.

c) A billiard ball of mass m traveling in the x-direction with velocity v_{1o} collides elastically with a second billiard ball of equal mass that is initially at rest. This time the two balls fly off with an angle between them. What is the angle? With some work one can solve the momentum conservation equations for both the x- and y-components of the momentum along with the kinetic energy conservation equation (**Exercise**: Do it.).

But here is an excellent example of where drawing diagrams can save a lot of agony. Momentum is a vector. Hence momentum conservation demands that $\mathbf{p}_{1o} = \mathbf{p}_{1f} + \mathbf{p}_{2f}$, or since the m's cancel out, $\mathbf{v}_{1o} = \mathbf{v}_{1f} + \mathbf{v}_{2f}$. By the rules of vector addition from Chapter 1, these three vectors must form a triangle with long side $\mathbf{v}_{1o}$.



Further, since the collision is elastic and kinetic energy is conserved as in (a), $v_{1o}^2 = v_{1f}^2 + v_{2f}^2$, which we recognize as the Pythagorean theorem for a right triangle with sides v_{1f} and v_{2f}. Therefore, the angle between $\mathbf{v}_{1f}$ and $\mathbf{v}_{2f}$ must be $90°$, the angle made by the two balls flying apart.

d) An object of mass M traveling horizontally at a velocity of v_o to the right collides inelastically with an object of mass m initially at rest. Because this is an inelastic collision, the masses stick together and after the collision must be moving with the *same* velocity v_f. Consequently, momentum conservation demands:

$$Mv_o = (M + m)v_f$$

or,

$$v_f = \frac{M}{M + m} v_o \,.$$

Notice that since the final mass $(M + m)$ is greater than the initial mass M, the final velocity must be less than the initial velocity.

More importantly, note that kinetic energy after the collision is by definition $KE_f \equiv \frac{1}{2}m_f v_f^2$, which from above is

$$KE_f = \frac{1}{2}(M+m)\left(\frac{M}{M+m}\right)^2 v_o^2 = \frac{1}{2}\frac{M^2}{M+m}v_o^2 = \frac{M}{M+m}KE_o$$

If $M = m$, this implies that $KE_f = 1/2KE_o$. The potential energy has not changed; in fact, there is none in this problem. We must conclude, as stated above, that **energy is not conserved in an inelastic collision.** It is dissipated as heat.

We saw in example (a) that two particles of equal mass exchanged velocity in an elastic collision. In this case, because m_2 was at rest before the collision, the initial difference in speeds was just v_{1o} and because m_1 was at rest after the collision, the final difference in speeds was just $v_{2f} = v_{1o}$. Thus, the difference in speeds, referred to as the *relative speed*, remained unchanged. This important result is true for all elastic collisions regardless of their masses and initial velocities. **The relative speed between two particles remains unchanged during an elastic collision.**

Exercise: Using conservation of momentum and conservation of kinetic energy in a one-dimensional collision between two masses, m_1 and m_2, show that the relative speed between the two particles is preserved. (For solution, see problem 3.74.)

13. Center of Mass and Momentum

Closely related to the concept of momentum is that of **center of mass**, usually abbreviated CM. In the previous sections we have tacitly assumed that the various objects behave like point particles whose entire masses are concentrated at their centers. For a single, symmetrical object, such as a block or billiard ball, the center of the object *is* the center of mass. All the velocities and accelerations that have figured throughout have in fact referred to the CM of the moving objects.

For a collection of masses, the situation is more complex. You know that to balance a seesaw, the heavier person must sit closer to the fulcrum than the lighter person. The balance point, or the position of the fulcrum, is the center of mass of the system.

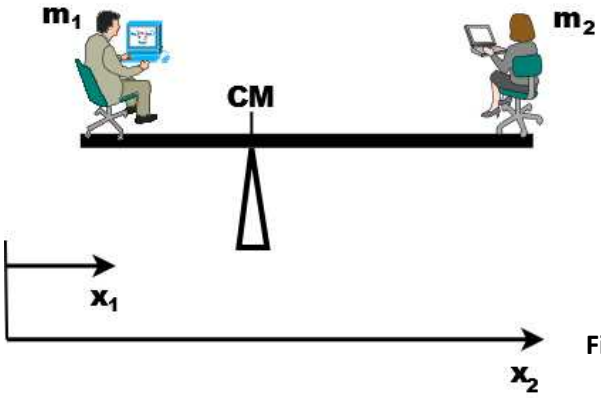


Figure 3.12

Mathematically, the position of the center of mass is defined as the "average" position of all the objects measured from any convenient origin:

$$x_{CM} = \frac{m_1x_1 + m_2x_2 + \cdots + m_nx_n}{m_1 + m_2 + \cdots + m_n} \quad (3.18)$$

where the denominator is M_{tot}, the system's total mass. For example, if you place the origin at the fulcrum, then $x_{CM} = 0$, and if there are two people of equal mass sitting on the seesaw, Eq. (3.18) gives $x_1 = -x_2$, as you'd expect. (**Exercise:** If one person is three times the mass of the other, where is the CM?)

Now, Eq. (3.18) implies that the velocity of the system's center-of-mass must be

$$v_{CM} = \frac{m_1 v_1 + m_2 v_2 + \cdots + m_n v_n}{m_1 + m_2 + \cdots + m_n} \qquad (3.19)$$

But the numerator is just the sum of all the momenta of the objects: $p_1 + p_2 + \cdots + p_n$. If the system is isolated, we know that its total momentum cannot change. Eq. (3.19) then immediately implies

The velocity of the center of mass of an isolated system remains constant

In particular, if the center of mass of an isolated system is initially at rest, it remains at rest for eternity.

Example 3.11: Isolated box. Suppose we have a box of length L with a partition dividing the box in half, as shown in Figure 3.13. The left half of the box is filled with air and the right half contains a vacuum. Apart from the air itself the box is massless and rests on a frictionless surface. At time $t = 0$ the partition magically vanishes and the air rushes from the left half of the box to the right, filling the box uniformly. Does the box move? If so, how much and in what direction?

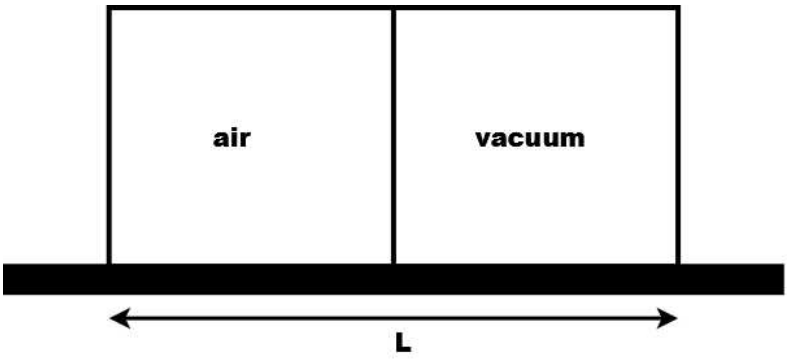


Figure 3.13

All one needs to solve this problem is momentum conservation and clean thought. The box on the frictionless surface represents an isolated system. It is initially at rest and so by the previous theorem, we know that its center of mass cannot move. Because the air represents the entire mass of the system, initially the CM is located at the position of the CM of the air itself—a distance $L/4$ from the left edge. After the partition vanishes, the air uniformly spreads throughout the box, and so the new CM must be at a distance $L/2$ from the left edge. But no outside force has been exerted on the box and so *the CM cannot have moved*. Therefore, the box must shift a distance $L/4$ to the left to keep the CM in its original position.

14. Impulse

Impulse is not a new concept; it is merely another word for the momentum imparted to an object during a short time period, typically during a collision. In other words, from the definition of force in Eq. (3.2), we know that $F = \Delta p/\Delta t$. Turning this around gives

$$\Delta p = F\Delta t \qquad (3.20)$$

If the force is constant over the interval Δt, for instance during a hammer blow, this allows us to calculate the change in momentum in a simple way. *The change of momentum is the impulse*.

Generally, though, the force is not constant. However, even in that case, the total change in momentum will be the total area under a force-vs.-time graph, or in analogy to the discussion in Section 2.3,

$$\Delta p = F_{avg}\Delta t \quad (3.21)$$

Example 3.12: Impulse. If one strikes an object such that the force is $F = 100$ N for 0.5 s, as shown in Figure 3.14a, then the impulse is $\Delta p = 50$ N s. If, on the other hand, the force rises as shown in Figure 3.14b, then the area is half that of (a) and so the impulse is only $\Delta p = 25$ N s.

Exercise: Suppose a 1000 kg car crashes into a wall at 30 m/s, such that the crash lasts 0.1 s. What is the impulse? Do you think anyone would survive without airbags?

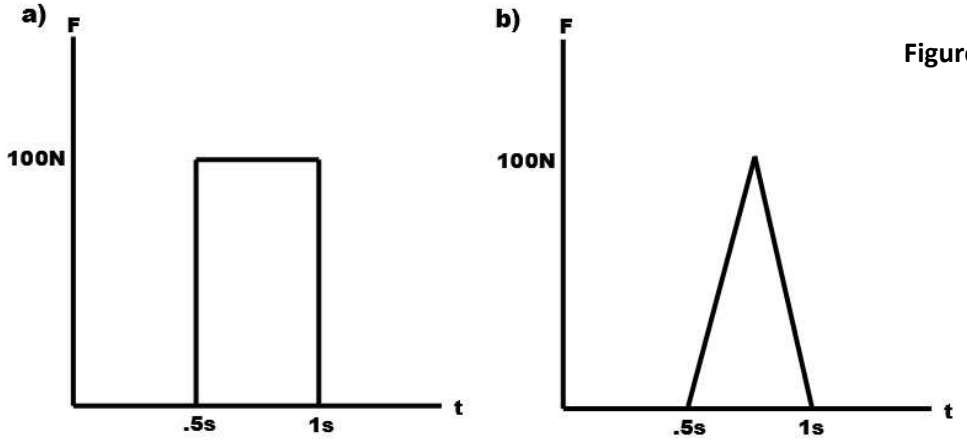


Figure 3.14

Example 3.13: Drag force and terminal velocity. This example is more advanced than the previous, but it is extremely useful for many problems in physics.

An object with a mass M and a cross sectional area A falls through the air. Find the terminal velocity v_T at which the object stops accelerating and begins to fall at a constant velocity.

The condition that the object reach terminal velocity is that the sum of the vertical forces acting on it equals zero. The only vertical forces acting on the object are gravity and air resistance, also known as the **drag force.** We thus require that

$$F_d = Mg,$$

for drag force F_d.

The entire problem therefore reduces to a calculation of F_d. Consider Figure 3.15 below. The object has cross sectional area A and at any instant of time is falling with velocity v. From the object's perspective, air molecules are rushing up to hit it with speed v in the opposite direction. Air molecules strike the object and either stick or bounce off.

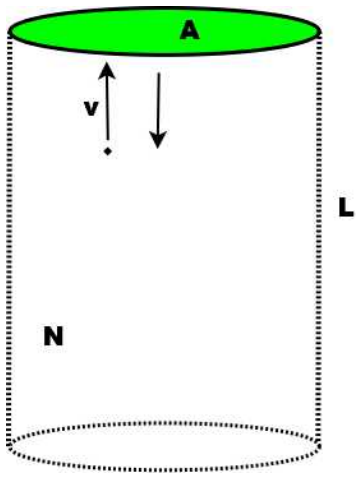


Figure 3.15

Eq. (3.20) tells us that $F = \Delta p/\Delta t$, which can be written

$$F = \frac{\Delta p}{\Delta t} = \frac{\Delta p}{collision} \times \frac{collisions}{second}.$$

If an air particle with mass m hits the object with velocity v and reverses direction in a perfectly elastic collision, we know from Example 3.10 that $\Delta p/collision = 2mv$. If the collision is perfectly inelastic, the air molecule sticks to the object and $\Delta p/collision = mv$. We can therefore write

$$\Delta p/collision = kmv,$$

where k is some number satisfying $1 \leq k \leq 2$. That should cover all collisions.

To calculate the number of collisions per second, consider the imaginary cylinder drawn in Figure 3.15. The cylinder has length L and contains N air molecules. Since from the perspective of the falling object the air molecules are traveling upward at velocity v, we can say that $L = vt$, where t is the time it takes a molecule at the bottom of the cylinder to reach the top, at which time all N molecules in the cylinder will have struck the object. The number of collisions per second is therefore just N/t. But $N = nV$, where n is the number density of air molecules and $V = AL$ is the volume of the cylinder. Thus, $N = ALn$, or $N = Anvt$. The number of collisions per second is then $N/t = Anv$ and the force is

$$F_d = kmv \times Anv = kmnAv^2$$

The quantity $mn = \rho_a$, the mass density of the air, and so finally

$$F_d = Dv^2 \quad (3.22)$$

$D \equiv k\rho_a A$ is called the **drag coefficient**. If you look up in textbooks drag coefficients for various objects you will discover that they are no more than what has just been derived. Usually the object is assumed to be spherical with cross section $A = \pi r^2$. The constant k must be measured, but $k \approx 1$.

The important thing about drag force is that if it is due to impulsive collisions, as we have assumed, $F_d \propto v^2$.

(At low velocities air can move smoothly over an object, in which case one finds that $F_d \propto v$. This is termed "laminar flow.")

The terminal velocity is now found by setting $Dv_T^2 = Mg$:

$$v_T = \sqrt{\frac{Mg}{D}} \quad (3.23)$$

Example 3.14: The bigger they come, the harder they fall. What is the terminal velocity for a human being? What is the force with which a human being falling at terminal velocity hits the ground compared to a mouse?

From Eq. (3.23), $v_T^2 = Mg/k\rho_a A$. Let us assume that the human being has mass $M = \ell A\rho$, where A is the falling person's cross sectional area (height × width), ℓ his or her depth (distance from front to back) and ρ the density of a human. Notice that the A's cancel out. We have already decided that $k \approx 1$. Then,

$$v_T^2 \approx \frac{\ell g \rho}{\rho_a}.$$

A good number for your number bank is that the density of air is $\rho_a \approx 1\ \text{kg/m}^3$. Also, from personal experience, you know that most, but not all, people float in water. This tells you that $\rho \leq 1000\ \text{kg/m}^3$. A prone person's depth ℓ must be about 25 cm. Plugging these numbers in yields, $v_T^2 \approx 2500$, or that $v_T \approx 50\ \text{m/s}$, in good agreement with the "online" number of 56 m/s. If this number had been a factor of ten larger, it would have been greater than the speed of sound and you would not have believed it. If it had been a factor of ten smaller, it would have meant that the person is practically floating, rather than falling, downward and you would not have believed it either. Thus, it passes the reality check.

To calculate the force of impact, we know $\Delta p/\Delta t = F$. On impact the velocity goes from v_T to zero, so $\Delta p = Mv_T$. The time it takes to come to rest must be $\Delta t \approx \ell/v_T$, the time necessary to cross the depth of the body ℓ at a velocity v_T. Thus,

$$F \approx \frac{Mv_T^2}{\ell} \quad (3.24)\,.$$

With the above expressions for M and v_T^2, this becomes

$$F \approx \frac{\ell A \rho^2 g}{\rho_a}.$$

Note that ℓA is the volume V of the falling object. Assuming that a human being and a mouse have the same density, the ratio of the force of impact for a human being to that of a mouse is just

$$\frac{F_h}{F_m} \approx \frac{V_h}{V_m}$$

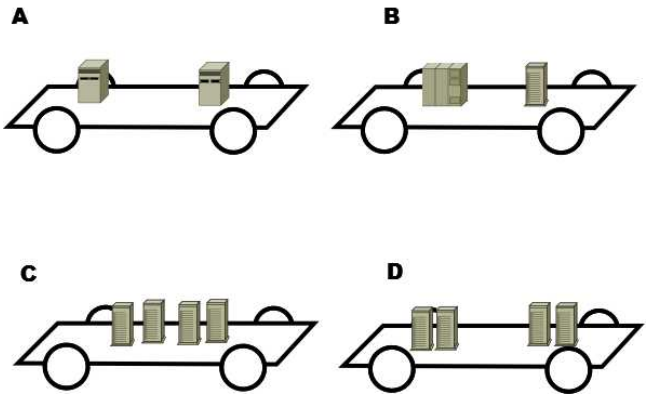
If the mouse's dimensions are $10 \times 5 \times 5\ \text{cm}^3$ and the human's $2 \times .5 \times .25\ \text{m}^3$, then $V_h/V_m = 1000$. Indeed, the bigger they come, the harder they fall. With $M = 80\ \text{kg}$ for a human, $v_T = 50\ \text{m/s}$, $\ell = 0.25\ \text{m}$, and $F_h = 8 \times 10^5\ \text{N}$, or 1000 gees, hardly survivable. On the other hand, the mouse experiences 800 N, or 250 gees. However, as an exercise you can calculate that the crash lasts only about 2 milliseconds, and so the mouse could survive, with severe injuries.

Note that you could have guessed the main result, Eq. (3.24). By the work-energy theorem the work involved in the crash is $W = Fd = 1/2Mv^2$, where d is the distance over which the force of impact operates, equal to ℓ above. Thus, the force of impact should be $F = Mv_T^2/2d$, which is one-half the above estimate. If we had used the average velocity during the crash, $v_T/2$, we would have gotten exactly the same answer. All roads lead to Rome.

PARTICLE MECHANICS PROBLEMS

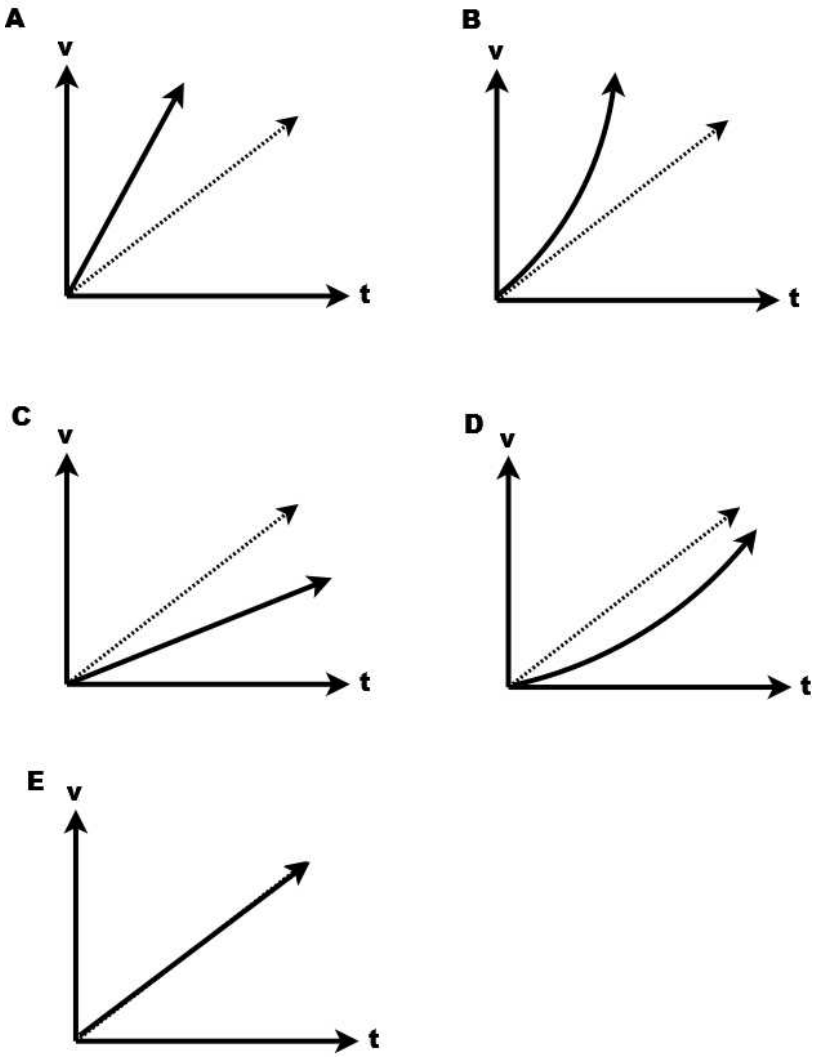
Problems 1-40 are at the subject test level. Thereafter, the problems are at the AP level.

1. Four identical carts are each loaded with computers distributed differently, as shown to the right, but loaded such that the total mass is the same. Identical forces are applied to the left edge of the carts for a period of 12 seconds. They are rolling over identical surfaces. At the end of the 12 seconds, which car will be moving with the smallest velocity?



(A) A
(B) B and D
(C) C
(D) D
(E) They will all be moving with the same velocity.

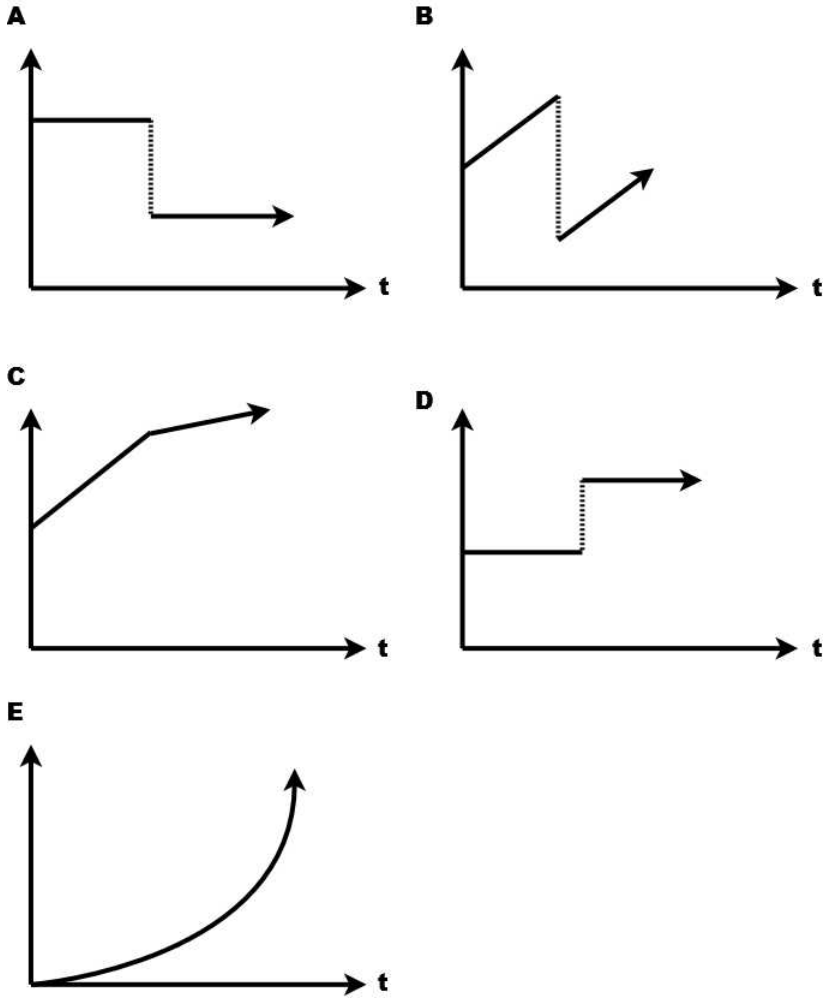
2. Suppose you jump on cart A in the previous problem just at the instant it begins moving. A possible velocity-vs.-time graph of the cart would then be which of the following? The dotted line gives the previous velocity.



(A) A
(B) B
(C) C
(D) D
(E) E (same velocity as previously)

3. You are pulling a block of mass m across a frictionless surface with a force F. At a certain place the surface becomes rough and begins to exert a force of friction on the block, but you continue to pull the block with the same force as before. Which of the below diagrams could represent the acceleration and velocity, respectively?

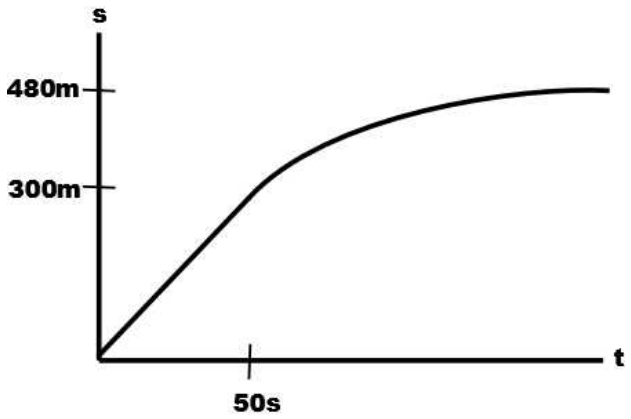
(A) A and B
(B) C and E
(C) A and C
(D) D and E
(E) A and E



4. A circus cannon fires an acrobatic dog, Astro, of mass m, into a net. The speed of Astro as he leaves the cannon is v. A cat, Beta, of mass $0.5m$ is then fired from the cannon. Assuming that the force exerted on each is constant throughout the cannon barrel, what is the speed of Beta as she leaves the mouth of the cannon?

(A) $2v$
(B) $\sqrt{2}v$
(C) v
(D) $v/\sqrt{2}$
(E) $v/2$

Questions 5 - 9 refer to the following figure.



5. A cyclist whose mass including the bicycle is 90 kg rides along a straight road for 50 s, as indicated in the figure. He then stops pedaling and slowly coasts to a stop. His average deceleration while stopping is nearest:

   (A) $0.01\ \mathrm{m/s^2}$
   (B) $0.05\ \mathrm{m/s^2}$
   (C) $0.1\ m/s^2$
   (D) $0.5\ \mathrm{m/s^2}$
   (E) $1\ \mathrm{m/s^2}$

6. Ten seconds after the cyclist has begun to slow down he has traveled an additional distance of:

   (A) 5.5 m
   (B) 10 m
   (C) 25 m
   (D) 55 m
   (E) 100 m

7. What is the average force applied to the cyclist and bicycle over his stopping distance?

   (A) 900 N
   (B) 90 N
   (C) 9 N
   (D) 0.9 N
   (E) 0 N

8. After how many seconds does the cyclist come to a stop?

   (A) 10 s
   (B) 20 s
   (C) 30 s
   (D) 40 s
   (E) 60 s

9. How much work does the external force perform in stopping the rider and bike?

(A) 900 N
(B) 1620 J
(C) 50 N
(D) 2000 J
(E) 90 J

10. An object moving at a constant velocity is acted on by several forces. Which of the following is true?

i. All the forces are acting in the same or opposite directions.
ii. The vector sum of the forces must be zero.
iii. The magnitude of the forces must be the same.

(A) i
(B) ii
(C) iii
(D) i and iii
(E) i, ii and iii

11. You are sitting on a swing that is tied to a branch of a big tree above you. Consider the following forces:

i. The force of gravity the Earth exerts on you
ii. The force of gravity you exert on the Earth
iii. The force the rope exerts on the swing
iv. The force the swing exerts on the rope

Which are action-reaction pairs?

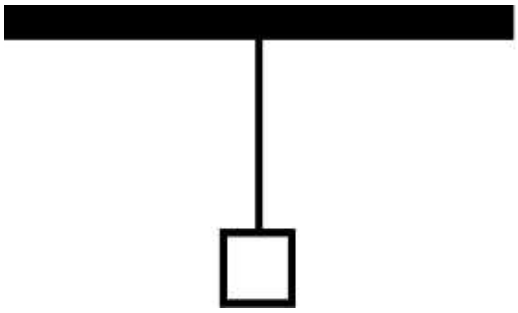
(A) Pair i and ii
(B) Pair iii and iv
(C) Pair i and iii
(D) Pair ii and iv
(E) Both pairs i and ii and iii and iv

12. You are driving along in your car and encounter a truck with a dead battery. Being a good Samaritan you decide to give the truck a push, as shown below.

The mass of your car is 1000 kg and the mass of the truck is 3000 kg. Your car wheels exert a force of 5000 N against the ground. What is the force your car exerts against the truck?

(A) 1250 N
(B) 5000/3 N
(C) 3750 N
(D) 5000 N
(E) 15,000 N

13. A block of mass 5 kg hangs by a string from the ceiling.



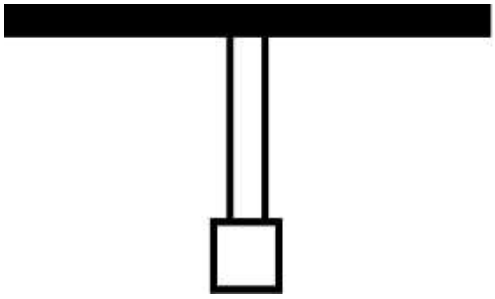
The force the ceiling exerts on the string is most nearly:

(A) 1 N, acting up
(B) 5 N, acting down
(C) 50 N, acting down
(D) 50 N, acting up
(E) 100 N, acting up

14. The string in the previous problem can support a maximum tension of 75 N before breaking. Several cats take turns jumping onto the block. The cat with the minimum mass necessary to break the string has a mass nearest:

(A) 10 kg
(B) 7.5 kg
(C) 5 kg
(D) 2.5 kg
(E) 1.5 kg

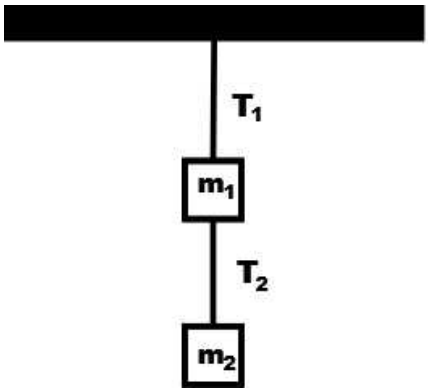
15. Say the single string in problem 13 has a tension T while supporting a 5 kg block. The single string is replaced by two strings, as shown.



The tension in each string is now:

(A) $0.25T$
(B) $0.5T$
(C) T
(D) $2T$
(E) $2.5T$

Questions 16 - 17 refer to the following figure.

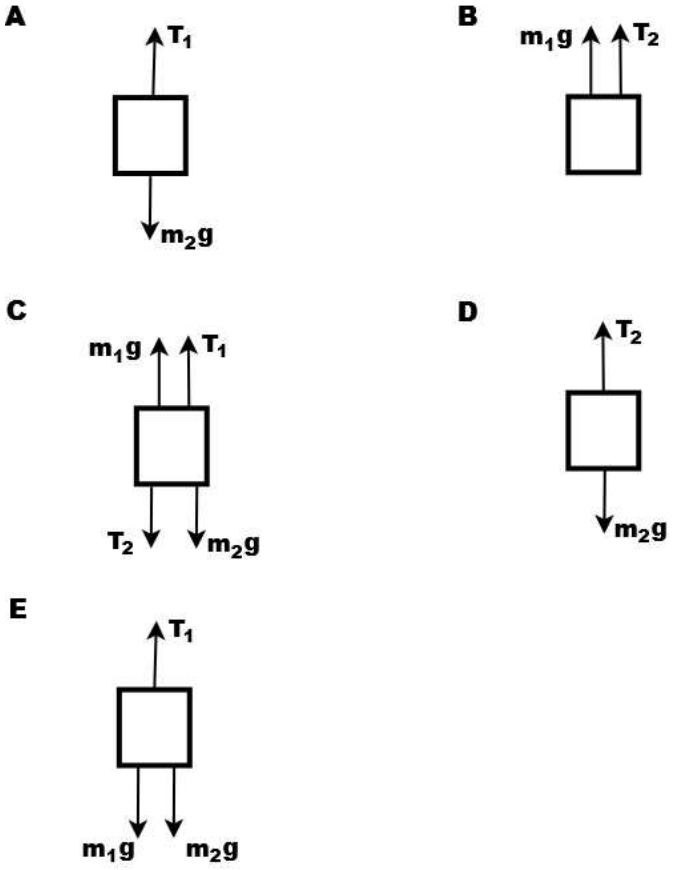


16. Two blocks of mass m_1 and m_2 are hanging by light strings from the ceiling, as shown, with the indicated tensions T_1 and T_2. Which of the following statements are true?

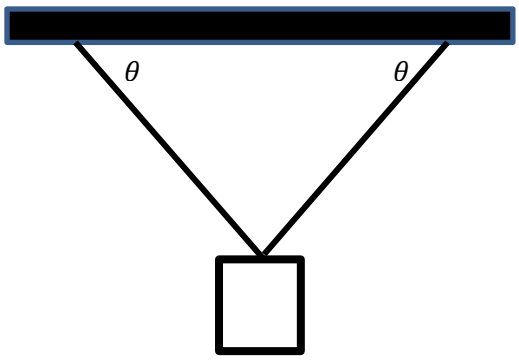
(A) T_2 is always greater than T_1
(B) T_2 is always less than T_1
(C) $T_1 = m_1 g + m_2 g$
(D) $T_2 = T_1 - m_1 g$
(E) (B), (C) and (D)

17. Which pair of the below diagrams correctly describes the forces on m_1 and m_2, respectively?

(A) A and B
(B) B and D
(C) B and D
(D) D and E
(E) E and D



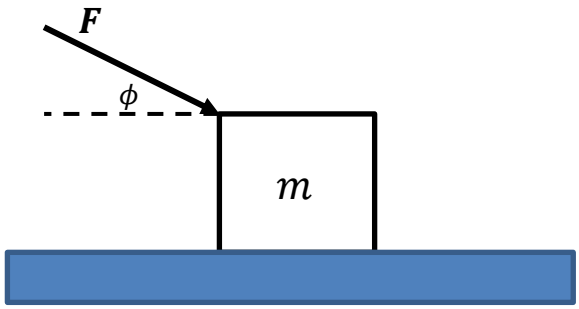
18. A block of mass m is supported by two strings, each making an angle θ between 0° and 90° with the ceiling, as shown.



The tension in each string will exceed the weight of the block

(A) when θ is less than 30°
(B) when θ is less than 45°
(C) when θ is greater than 30°
(D) always
(E) never

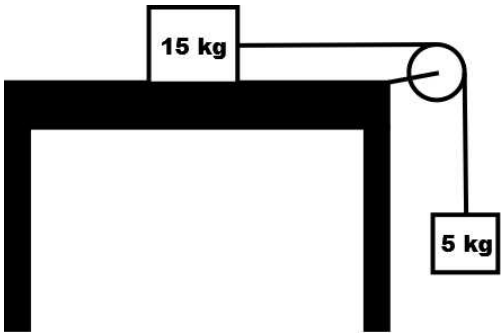
19. A force $\boldsymbol{F}$ is applied to a mass m at an angle ϕ above the horizontal, as shown below.



The mass rests on a frictionless table. The magnitude of the acceleration, a, and the magnitude of the normal force, N, are given by which of the following?

(A) $a = F\cos\phi/m$; $N = F\sin\phi + mg$
(B) $a = F/m$; $N = mg$
(C) $a = F\cos\phi/m$; $N = mg$.
(D) $a = F\cos\phi/m$; $N = F\sin\phi$
(E) $a = F/m$; $N = F\sin\phi$

Questions 20 - 21 refer to the following figure.



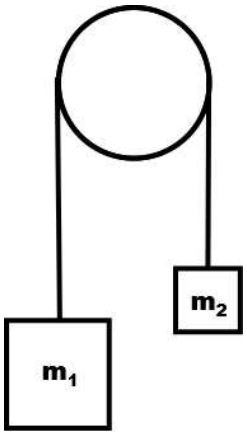
20. Two masses, one of 5 kg and one of 15 kg, are connected by a string of negligible mass, as shown in the figure. The string passes over a pulley, also of negligible mass, and the 5 kg mass hangs off the edge of a table. Assuming the table is frictionless, what is the approximate acceleration of the 15 kg mass?

    (A) $10\ \mathrm{m/s^2}$
    (B) $7.5\ \mathrm{m/s^2}$
    (C) $5\ \mathrm{m/s^2}$
    (D) $2.5\ \mathrm{m/s^2}$
    (E) $0\ \mathrm{m/s^2}$

21. Consider the same setup as in problem 20, except that now friction exists between the 15 kg block and the table, with a coefficient of kinetic friction μ. Consider the new tension T in the string and the new acceleration of the blocks, a. One can conclude:

    (A) T and a are both the same as they were in the case without friction.
    (B) T and a are both larger than they were in the case without friction.
    (C) T and a are both smaller than they were in the case without friction.
    (D) T is larger than it was in the case without friction, but a is smaller.
    (E) T is smaller than it was in the case without friction, but a is larger.

Questions 22 - 23 refer to the following figure.



22. Two masses, m_1 and m_2, are connected by a string of negligible mass, which passes over a massless pulley, as shown in the figure. Assume $m_1 > m_2$. What is the acceleration a of m_1?

(A) $a = (m_1 - m_2)g/(m_1 + m_2)$
(B) $a = (m_1 + m_2)g/(m_1 - m_2)$
(C) $a = (m_1 m_2)g/(m_1 + m_2)$
(D) $a = (m_2 - m_1)g/(m_1 + m_2)$
(E) $a = (m_1 - m_2)g/(m_1 m_2)$

23. Assume the same setup as in the previous problem, with $m_1 = 8$ kg and $m_2 = 4$ kg. Considering the tensions in the string to the left and right of the pulley:

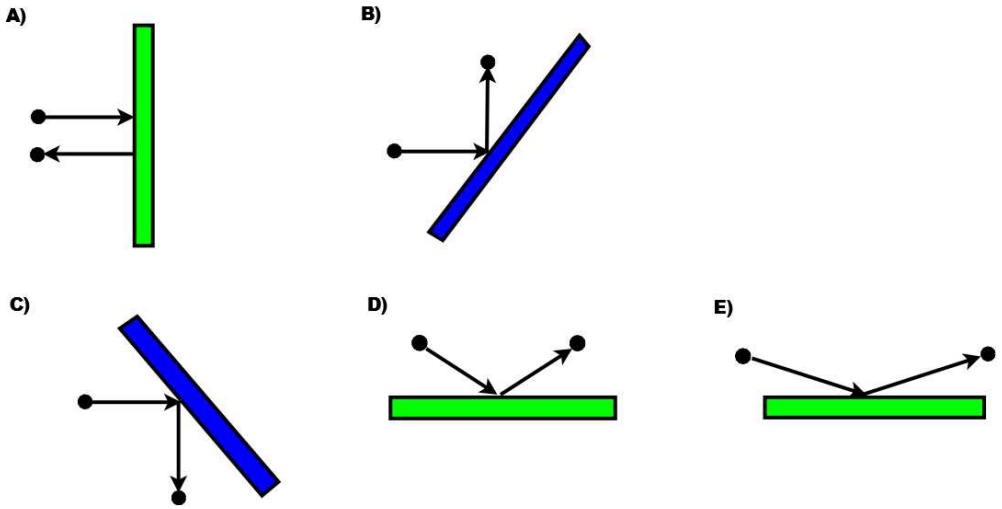
(A) The tension on the right is less than the tension on the left.
(B) The two tensions are equal.
(C) The tension on the left is less than 80 N.
(D) The tension on the right is greater than 40 N.
(E) All of the above except (A)

24. The units of momentum can be expressed as

(A) $\mathrm{N \cdot s}$
(B) $\sqrt{\mathrm{kg \cdot J}}$
(C) $\sqrt{\mathrm{kg \cdot W \cdot s}}$
(D) all of the above
(E) none of the above

Questions 25 - 28 refer to the following information.

Shown in the figure are several elastic collisions between a ball and a surface. Assume the x-direction is horizontal and the y-direction is vertical. The ball has mass m and initial speed v.



25. In which diagram is the change in the ball's x-momentum the greatest?

    (A) A
    (B) B
    (C) C
    (D) D
    (E) E

26. In which diagram does the ball's y-momentum change by $-mv$?

    (A) A
    (B) B
    (C) C
    (D) D
    (E) E

27. For what angle ϕ from the horizontal in diagram D will the change in y-momentum equal the change in diagram B?

    (A) 90°
    (B) 60°
    (C) 45°
    (D) 30°
    (E) None of the above

28. Again assuming an elastic collision between the ball and the surface, the change in kinetic energy during the collision

    (A) is greatest in diagram A because the ball's velocity reverses.
    (B) is the same nonzero value for all of them.
    (C) is zero for all of them.
    (D) is least in diagram E, since the ball's velocity changes direction least.
    (E) is positive for B and negative for C.

29. A spring obeys Hooke's law. If an amount of work W is required to stretch the spring a length x beyond its unstretched length, how much work does it take to stretch it to $3x$?

    (A) W
    (B) $3W$
    (C) $6W$
    (D) $9W$
    (E) $27W$

30. Which of the following are units of power? (Pascal is a unit of pressure.)

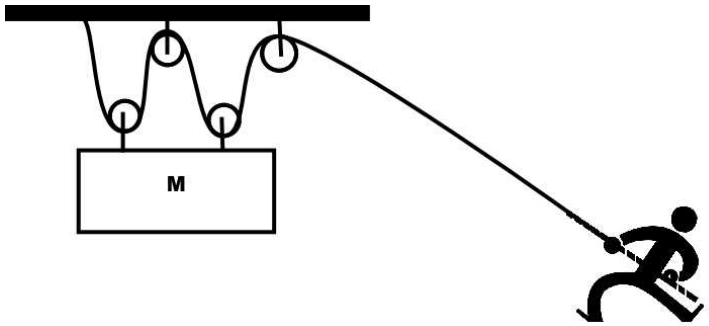
    i. watt ii. kilowatt hr iii. newton $\mathrm{m/s}$ iv. kilowatt $\mathrm{hr/day}$ v. pascal $\mathrm{m^3/s}$

    (A) i
    (B) i and ii
    (C) ii and iii
    (D) iv and v
    (E) i, iii, iv and v

31. A 20 kg chimp climbs three meters at a constant speed along a rope in 4 seconds. What is the power the chimp expends against gravity?

    (A) 100 W
    (B) 150 W
    (C) 200 W
    (D) 300 W
    (E) 600 W

Questions 32 - 35 refer to the following diagram.



32. Your father decides to lift a freezer into a truck by a pulley system, as shown. The mass of the freezer is $M = 100$ kg. Assume the masses of the rope and pulleys are negligible. To raise the freezer, he must exert a force most nearly equal to

    (A) 50 N
    (B) 250 N
    (C) 500 N
    (D) 750 N
    (E) 1000 N

33. If your father raises the freezer 2 m, its energy changes by

(A) 100 J
(B) 500 J
(C) 1000 J
(D) 1500 J
(E) 2000 J

34. If your father pulls on the rope at 0.4 m/s, how fast does the freezer rise?

(A) 0.1 m/s
(B) 0.2 m/s
(C) 0.25 m/s
(D) 0.4 m/s
(E) 0.8 m/s

35. Say it takes 20 s to raise the freezer 2 m. Which statement about the power is true?

(A) Your father expends 1/4 the power of the pulley system.
(B) Your father expends 4 times the power of the pulley system.
(C) Your father and the pulley system expend the same amount of power.
(D) The pulley system expends 1/4 the power as your father.
(E) The pulley system expends 4 times the power as your father.

36. Your dog Astro is nosing a dinner plate of mass M across a frozen pond at a constant velocity v. There is a coefficient of friction μ between the ice and the block. What is the rate of work Astro does on the plate?

(A) $\mu M v^2$
(B) $\mu M g v$
(C) $M g v^2/\mu$
(D) $\mu M v^2/g$
(E) $\mu M g/v$

37. You find yourself stranded on planet Alpha, where the gravitational field is 5 N/kg. Which of the following statements are true? The 10 kg mass you hold in your hand

i. will weigh about half as much on Alpha as on Earth.
ii. will weigh about twice as much on Alpha as on Earth.
iii. will have the same mass on Alpha as on Earth.
iv. will fall with a greater acceleration than on Earth.
v. will require a lesser force to push it across a horizontal surface with a small amount of friction.

(A) i and iv
(B) i, iii and v
(C) i, iii, iv and v
(D) ii, and iii
(E) ii, iii and iv

38. You find yourself stranded on planet Alpha, which is half as dense as Earth but which has a radius three times that of Earth's. What is your weight on Alpha compared to your weight on Earth?

    (A) $2/3$
    (B) The same
    (C) $3/2$
    (D) 3
    (E) 6

39. Two identical planets, Alpha and Beta, touch each other. Khan the Conqueror suddenly halves the density of Beta, and doubles the radius of Alpha. The gravitational force between them:

    (A) stays the same.
    (B) doubles.
    (C) halves.
    (D) goes up by a factor of $16/9$.
    (E) goes down by a factor of $9/16$.

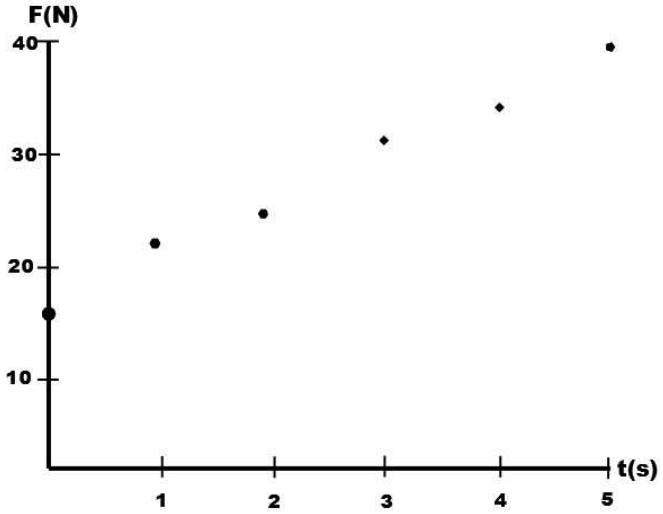
40. A 1 kg mass strikes a 2 kg mass in an inelastic collision on a horizontal frictionless surface. The 2 kg mass is initially at rest. After the collision the energy is

    (A) $1/3$ what is was before the collision.
    (B) $1/2$ what it was before the collision.
    (C) the same as it was before the collision.
    (D) twice what it was before the collision.
    (E) three times what it was before the collision.

41. An object of mass m_1 and an object of mass m_2 are initially attached to each other and at rest. At the instant $t = 0$ they are pushed apart by a spring that has been compressed between them. After $t = 0$, $v_1 = 2$ m/s and $v_2 = -3$ m/s. One can conclude that

    (A) $m_1/m_2 = 3/2$
    (B) $m_1 = 2$ kg and $m_2 = 3$ kg
    (C) $m_1 = 3$ kg and $m_2 = 2$ kg
    (D) $m_1/m_2 = 2/3$
    (E) $m_1/m_2 = -3/2$

Questions 42 - 43 refer to the following figure.



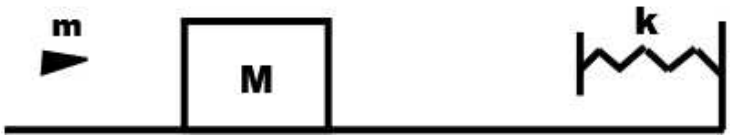
42. A student measures the force exerted on an object vs. time and plots the graph above. The change in force per second is most closely given by

(A) 15 N/s
(B) 12 N/s
(C) 10 N/s
(D) 8 N/s
(E) 5 N/s

43. The change in momentum of the object during the period covered by the graph is most closely

(A) 125 N s
(B) 100 N s
(C) 80 N s
(D) 50 N s
(E) 25 N s

44. As shown in the figure below, a bullet of mass m traveling with an initial velocity v_b hits a mass M, lodging itself in M. The combined system slides across a frictionless surface and strikes a spring with spring constant k, which compresses a distance d.



To determine v_b one needs:

(A) Conservation of energy
(B) Conservation of momentum
(C) M, m
(D) k, d
(E) All of the above

45. A $1\,\text{kg}$ mass strikes a $2\,\text{kg}$ mass in an inelastic collision on a horizontal frictionless surface. The $1\,\text{kg}$ mass is traveling at $1\,\text{m/s}$ in the positive direction and the $2\,\text{kg}$ mass is traveling at $1\,\text{m/s}$ in the negative direction. The final velocity of the masses is

    (A) $+1\,\text{m/s}$
    (B) $+1/2\,\text{m/s}$
    (C) $-1\,\text{m/s}$
    (D) $-2/3\,\text{m/s}$
    (E) $-1/3\,\text{m/s}$

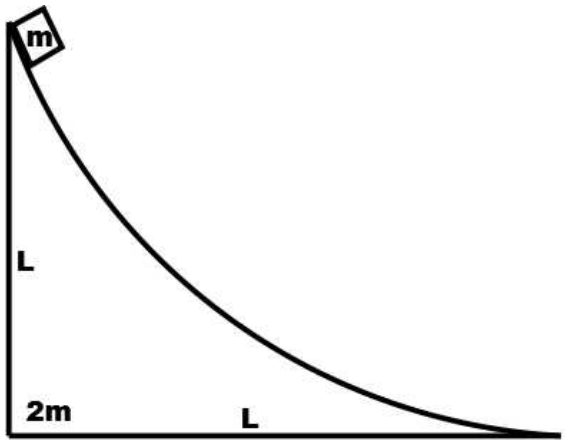
46. You find yourself stranded on planet Alpha again, which you remember is half as dense as Earth but which has a radius three times that of Earth's. What is the escape velocity v of Alpha compared to Earth's escape velocity?

    (A) $2/3$
    (B) The same
    (C) $3/2$
    (D) $3/\sqrt{2}$
    (E) 6

47. Two planets, Gamma and Delta, with masses m_G and m_D and radii R_G and R_D are initially at rest a large distance d apart from each other. Khan the Conqueror suddenly allows the two planets to approach under their mutual gravitational attraction until they collide. To determine the velocity at which Gamma and Delta collide you could

    (A) use conservation of energy alone.
    (B) use conservation of momentum alone.
    (C) use both conservation of energy and momentum.
    (D) The problem cannot be solved using energy and momentum conservation.
    (E) The problem cannot be solved unless you know how long it takes the two planets to collide.

Questions 48 - 49 refer to the following figure.



48. A crate of mass m sits atop a frictionless ramp that has a height and length L. The ramp has mass $2m$, which is concentrated at the lower left corner, and it sits on a frictionless ice pond. The crate is released and slides down the ramp. By the instant that the crate reaches the bottom, how far has the ramp moved and in what direction?

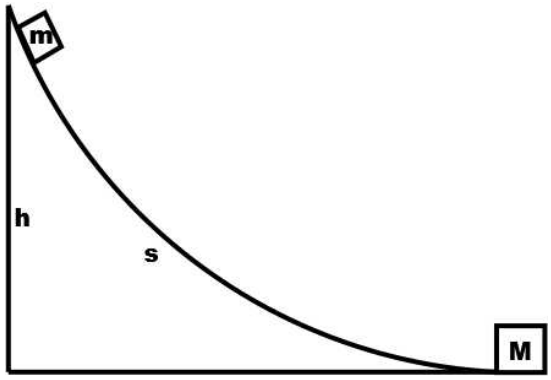
    (A) It hasn't moved.
    (B) It has moved to the left a distance $L/3$.
    (C) It has moved to the right a distance $L/3$.
    (D) It has moved to the left a distance $L/2$.
    (E) It has moved to the right a distance $L/2$.

49. If one wanted to calculate the velocity of the ramp as the box slid off, one would need to do which of the following?

    i. Equate the momentum of the box plus the momentum of the ramp to zero
    ii. Equate the kinetic energy of the crate to mgL
    iii. Equate the kinetic energy of the crate plus ramp to mgL

    a. i
    b. ii
    c. iii
    d. i and ii
    e. i and iii

Questions 50 - 52 refer to the following figure.



50. A crate of mass m is released from rest at the top of a stationary ramp of height h, whose length along the curve is s, as shown. Assuming that the ramp is frictionless, in calculating the speed of the crate v at the bottom of the ramp, one can assume:

    (A) $h = vt$
    (B) $mgh = 1/2\ mv^2$
    (C) $s = 1/2\ at^2$
    (D) $v^2 = 2as$
    (E) $h = 1/2\ gt^2$

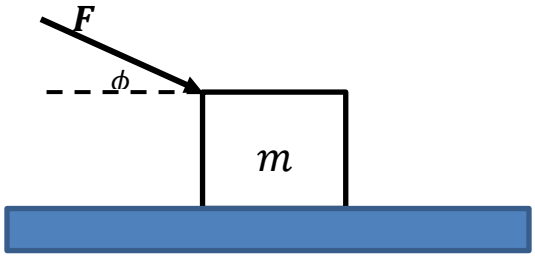
51. Suppose the crate of mass m in the figure collides with the crate of mass M in an inelastic collision at the bottom of the ramp. If $M = 2m$, then the final velocity of the system in terms of h will be

    (A) $v_f = \sqrt{2gh}/3$
    (B) $v_f = \sqrt{2gh}$
    (C) $v_f = 3\sqrt{gh}$
    (D) $v_f = gh/3$
    (E) $v_f = 2gh/3$

52. Suppose instead that the crate of mass m hits the crate of mass M in an elastic collision and that $M = m$. In this case one expects that

    (A) m will bounce off M and rise up to height h again.
    (B) m will bounce off M and rise up to a height less than h.
    (C) m will stop and M will move off with velocity $v_M = \sqrt{2gh}$.
    (D) both m and M will move off with velocity $v_M = \sqrt{2gh}$.
    (E) m will stop and M will also remain stationary.

53. Consider the same situation as in problem 19 except that there is now a coefficient of friction μ between the block and the table.



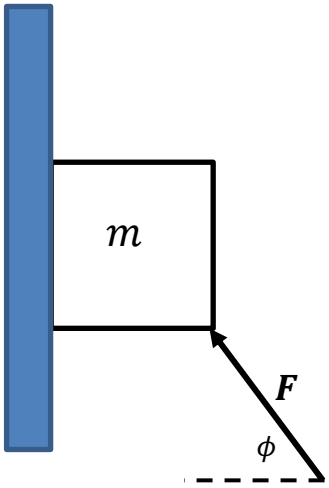
The acceleration is now

(A) $a = F(\cos\phi + \mu\sin\phi) + mg$
(B) $a = F\cos\phi/m + \mu mg$
(C) $a = F(\cos\phi - \mu\sin\phi)/m - \mu g$
(D) $a = (F\cos\phi + \mu\sin\phi)/m + \mu g$
(E) $a = (F + \mu\sin\phi)/m + g$

54. Consider the same situation as in the previous problem. Is there a critical angle less that 90° above which one cannot move the block no matter how great the force one applies?

(A) There is a critical angle that depends on both μ and g.
(B) There is a critical angle that depends on μ but does not depend on g.
(C) There is a critical angle that depends on g but does not depend on μ.
(D) There is a critical angle that depends on neither μ nor g.
(E) There is no critical angle.

Questions 55 - 56 refer to the following figure.



55. A block of mass m rests against a vertical wall, as shown. There is a nonzero coefficient of friction μ between the block and the wall. A force F is exerted against the corner of the block at an angle ϕ. To determine the magnitude of F sufficient to keep the block from sliding down the wall one needs to realize:

    i. The force of friction acts down; ii. The force of friction acts up; iii. The normal force N acts to the left; iv. There is no normal force in the problem; v. The normal force acts to the right.

    (A) i
    (B) i and iii
    (C) ii and iii
    (D) ii and iv
    (E) ii and v

56. The magnitude of **F** is

    (A) $F = mg/(\mu\cos\phi - \sin\phi)$
    (B) $F = mg/(\mu\cos\phi + \sin\phi)$
    (C) $F = g/(\mu\cos\phi + \sin\phi)$
    (D) $F = mg/(\cos\phi + \sin\phi)$
    (E) $F = mg/(\cos\phi + \mu\sin\phi)$

57. A 1 kg mass traveling in the positive x-direction at 3 m/s strikes a 3 kg mass traveling at 1 m/s in the negative x-direction. The collision is perfectly elastic. What are possible final velocities for the 1 kg mass and the 3 kg mass, respectively?

    (A) 0 m/s; 3 m/s
    (B) 3 m/s; −1 m/s
    (C) −3 m/s; 1 m/s
    (D) 3 m/s; 0 m/s
    (E) (B) and (C)

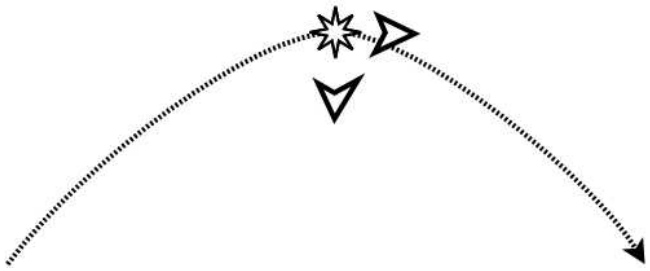
58. An open railroad car is passing at a constant speed v_o under a water tower, which suddenly dumps a lot of water into the car. The speed of the car will

    (A) stay the same because of momentum conservation.
    (B) increase because of momentum conservation.
    (C) decrease because of momentum conservation.
    (D) increase because of energy conservation.
    (E) stay the same because the water is falling in the y-direction and the car is moving in the x-direction.

59. Suppose in the previous problem that the amount of water dumped into the car is $0.20m$ over 5 seconds, where m is the mass of the car. The average external force required to keep the car moving at its original velocity is:

    (A) $4v_o m$
    (B) ma
    (C) $v_o m$
    (D) $.04v_o m$
    (E) zero

Questions 60 - 61 refer to the following figure.



60. A projectile is launched from the ground. The dotted line in the figure indicates the path the projectile would normally follow. At the top of its flight, however, when the projectile is moving horizontally, it explodes into two pieces. One piece falls straight down, as shown. At the instant of the explosion the second piece continues to move horizontally. Ignoring any air resistance, the second piece

    (A) follows the original trajectory.
    (B) hits the ground at a greater distance from the launch point with its launch speed.
    (C) hits the ground at a smaller distance from the launch point with its launch speed.
    (D) hits the ground at a greater distance from the launch point with a higher speed than its launch speed.
    (E) Hits the ground at a smaller distance from the launch point with a higher speed than its launch speed.

61. Consider the same scenario as in the previous problem. If R is the range of the projectile in the absence of an explosion and the projectile fragments into two equal masses, then the total distance fragment 2 travels is:

    (A) $2R/3$
    (B) R
    (C) $3R/2$
    (D) $2R$
    (E) $4R$

62. Pick all correct responses. Air bags in cars protect occupants in a crash by

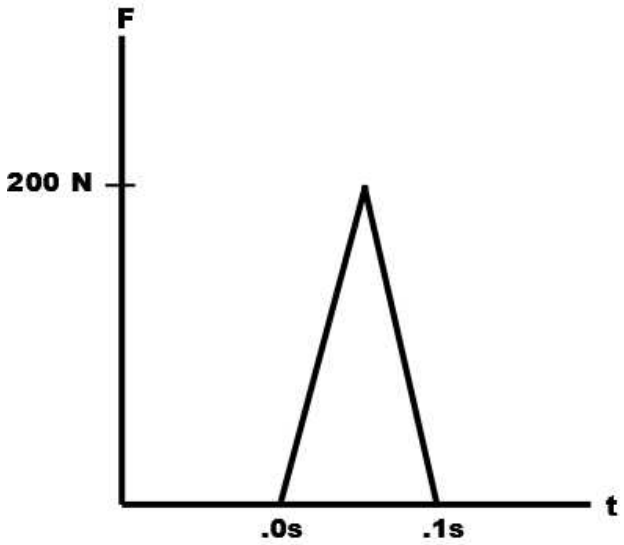
    (A) increasing the time needed for the occupant to come to a stop, thereby decreasing the force of impact.
    (B) increasing the momentum of the occupant.
    (C) decreasing the impulse of the occupant.
    (D) decreasing the amount of kinetic energy lost during the collision.
    (E) None of the above

63. Model rocket engines burn for only a short time before using up all their propellant. Suppose a model rocket is launched from its stand at an angle $60°$ above the horizontal. The mass of the rocket is 0.25 kg and its final speed is 50 m/s. If the engine burns for 1.25 s, the impulse it gives to the rocket is most nearly

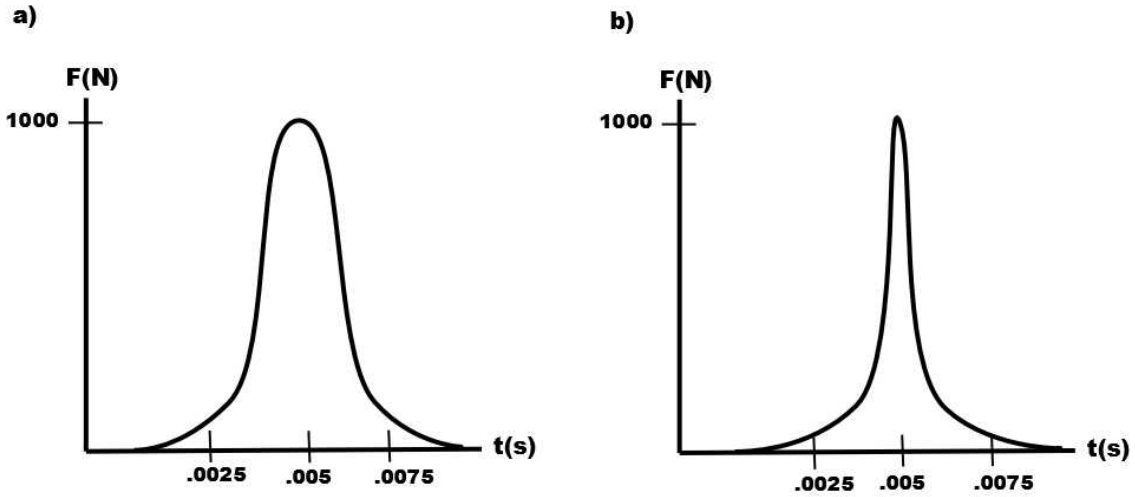
    (A) 15.6 N s
    (B) 12.5 N s
    (C) 6.25 N s
    (D) $\frac{\sqrt{3}}{2} \times 12.5$ N s
    (E) 12.5 N

64. A carnival "high striker" game is a game in which you hit a lever with a sledgehammer, trying to make a puck on the other end of the lever rise as high as possible. Suppose you are able to swing a sledgehammer such that the peak force on the lever is 200 N in the profile shown below and suppose that the puck has a mass of 1 kg. Ignoring friction and the lever, approximately how high will the puck rise in the tower?

(A) 20 m
(B) 15 m
(C) 10 m
(D) 5 m
(E) 1.25 m



65. A tennis ball of mass $m = 60$ g (slightly over regulation) hits a wall horizontally at 35 m/s and rebounds at 15 m/s. Which of the below diagrams might represent a graph of the force vs. time with which the ball hits the wall.



(A) (a)
(B) (b)
(C) Both (a) and (b)
(D) Neither (a) nor (b)
(E) Not enough information to draw a conclusion

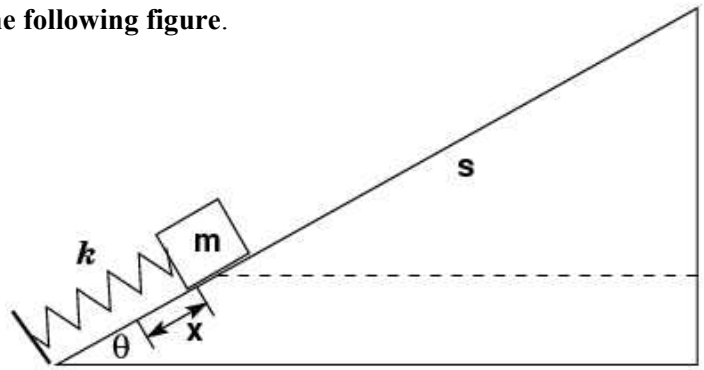
66. A 60 kg passenger is traveling in a car at 30 m/s when suddenly a tree steps onto the road. The car crashes and the air bags deploy. If the passenger must be stopped within 1 meter, the force on the passenger is nearly

(A) 270 N
(B) 1800 N
(C) 2700 N
(D) 18,000 N
(E) 27,000 N

67. Assuming constant deceleration, estimate the stopping time in the previous problem.

(A) 1/100 s
(B) 1/60 s
(C) 1/30 s
(D) 1/15 s
(E) 1 s

Questions 68 - 69 refer to the following figure.



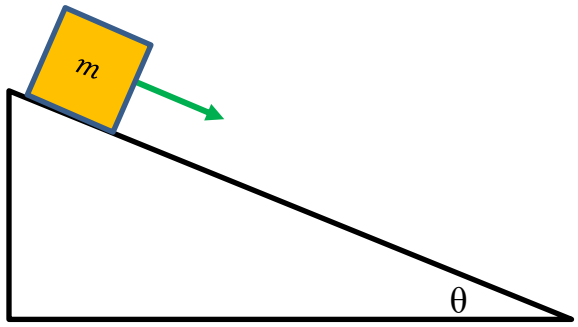
68. A box of mass m is released from rest at the top of a frictionless ramp tilted at an angle θ from the horizontal. The box slides down a distance s and collides with spring with a spring constant k. The spring then compresses a distance x. To find x one should

(A) equate mgs to $1/2\ kx^2$ and solve for x.
(B) equate $mgs\cos\theta$ to $1/2\ kx^2$ and solve for x.
(C) equate $mgs\sin\theta$ to $1/2\ kx^2$ and solve for x.
(D) equate $mg(x+s)\sin\theta$ to $1/2\ kx^2$ and solve for x.
(E) equate mg to kx and solve for x.

69. The maximum velocity of the box occurs at the instant when

(A) the box hits the spring.
(B) the potential energy is minimum.
(C) the kinetic energy is maximum.
(D) the forces on the block are zero.
(E) All of the above except (A)

Questions 70 - 73 refer to the following figure.



70. Two students, Alice and Bob, decide to compute the power that the Earth's gravitational field expends on a block of mass m as the block slides down a frictionless inclined plane.

Alice reasons: "The gravitational force pulling the block down the incline is $F = mg\sin\theta$. The block's velocity at any given height h from the top of the incline is $v = \sqrt{2gh}$. Power is defined as force × velocity. Therefore, the power is $P = mg\sin\theta\sqrt{2gh}$."

Bob reasons: "Power is $\Delta W/\Delta t$. By the work-energy theorem, the change in work is the change in kinetic energy, but without friction $\Delta W = mgh$. The change in time is $\Delta t = \Delta v/a$, $\Delta v = \sqrt{2gh}$ and $a = g\sin\theta$. So $\Delta t = \sqrt{2gh}\,/(g\sin\theta)$. Therefore, $P = mgh \times \frac{g\sin\theta}{\sqrt{2gh}} = mg\sin\theta\sqrt{\frac{gh}{2}}$."

Alice and Bob look at each other and scratch their heads. Who is correct?

i. Alice ii. Bob iii. Neither iv. Both v. The problem is imprecisely worded.

(A) i
(B) ii
(C) iii
(D) iv
(E) iv and v

71. Suppose that the incline in the above problem is tilted at $\theta = 30°$ and the coefficient of friction between the ramp and the block is $\mu = 0.2$. The mass of the block is 5 kg. After the block slides down the incline a vertical distance $h = 1$ m, what is the total work done by the force of friction?

(A) $10\sqrt{2}$ J
(B) $10\sqrt{3}$ J
(C) 50 J
(D) 75 J
(E) 100 J

72. With the information given in problem 71, what is the approximate instantaneous power expended by friction after the block has descended a vertical distance of $h = 1$ m? (Calculator is permissible.)

(A) +10 W
(B) +5 W
(C) −10 W
(D) −30 W
(E) −50 W

73. With the information given in problem 71, what is the approximate average power expended by friction as the block descends a vertical distance of $h = 1$ m? (Calculator is permissible.)

(A) $+15$ W
(B) $+10$ W
(C) -10 W
(D) -15 W
(E) -30 W

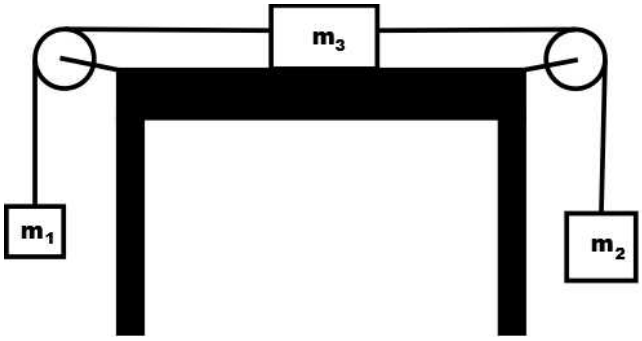
Free response questions

74. Two objects, m and M, with initial velocities v_o and U_o respectively, undergo a one-dimensional elastic collision. Show that their relative speed is unchanged during the collision. (The relative velocity between m and M is defined as $v - U$.)

75. Two masses, m_1 and m_2, are attached to each other by a massless string, which passes over two pulleys of negligible mass, as shown in the figure. Assume that $m_2 > m_1$.

a) Draw below all the forces on m_1 and m_2. Draw accurately the lengths of the vectors to indicate the relative magnitude of the forces.

b) Find the acceleration of m_1. Is it greater than, equal to, or less than the acceleration of m_2?



A third mass, m_3, is added to the system as shown. Say $m_1 = 5$ kg, $m_2 = 10$ kg and $m_3 = 15$ kg.

c) Draw the forces on the three blocks. As before, draw accurately the length of the vectors to indicate the relative size of the forces.

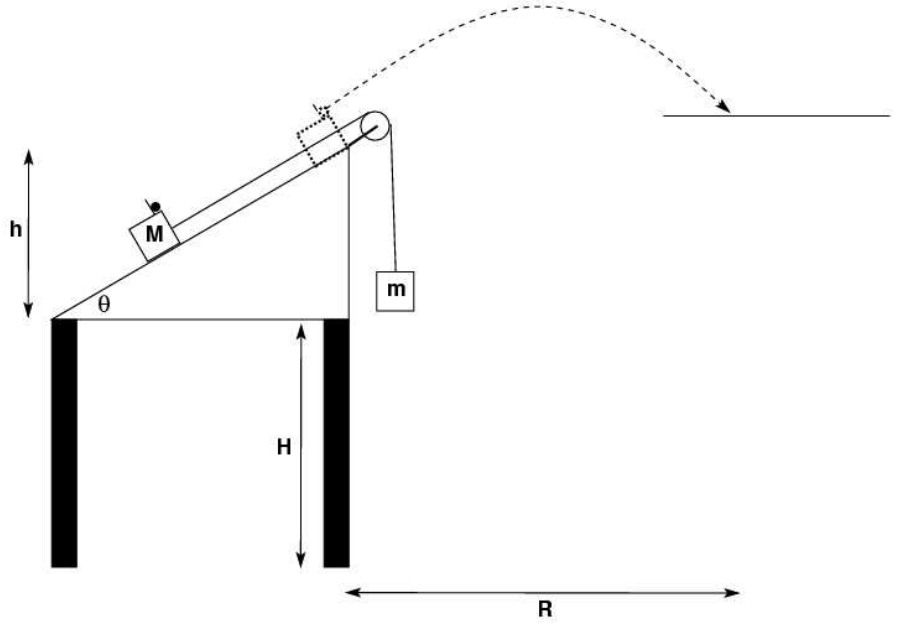
d) Is the acceleration now greater, less than, or equal to what it was before? Justify your answer.

e) Calculate the acceleration of m_3.

f) Is the tension in the left-hand string greater than, equal to or less than the tension in the right-hand string?

g) Assuming $g = 10\ \text{m/s}^2$, calculate the tension in the two strings.

76. A mass m is attached to a mass M by a light string that passes over a frictionless pulley of negligible mass, so that it can pull M up the incline. The incline is frictionless and tilted at an angle of $\theta = 30°$ from the horizontal. M is initially at rest at the bottom of the incline; then m is released. When M reaches the top it hits a stop and launches a small ball of negligible mass from a height $h = 1$ m above the starting position. The ball lands on a shelf at the same height at a range $R = 1.8\sqrt{3}$ m from its launch point. Assume the acceleration of gravity is $g = 10\ \mathrm{m/s^2}$.



a) What is the launch speed of the ball?

b) What is the acceleration of M?

c) What is the mass M in terms of m?

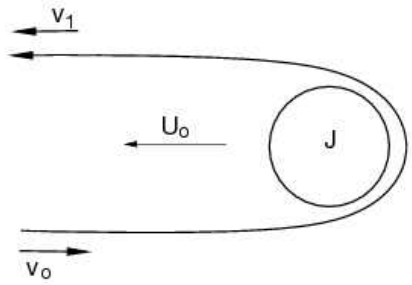
d) If the ball is given a larger mass, will R increase, decrease, or stay the same? Explain your reasoning.

e) The shelf is removed and the experiment repeated. This time the ball falls to the floor. If $H = 1$ m, what is the time the ball spends in the air?

f) What is the new range, R?

77. Superman is able to leap over tall buildings in a single bound. The tallest building in the world is currently the Burj Khalifa in Dubai, which is about 800 m tall and will soon be overtaken by the Jeddah Tower in Saudia Arabia, which will be 1 km tall. Assuming that Superman's legs act like springs and that he can bend them 0.5 m, estimate the strength of Krypton's gravitational acceleration relative to the Earth's.

78. In an attempt to save millions of taxpayer dollars, NASA frequently uses the gravitational field of a planet, in particular Jupiter, to give space probes a "gravitational assist" along their trajectories to the outer solar system. Assume that this "gravitational slingshot" can be modelled as a perfectly elastic collision, that Jupiter has an orbital velocity of U_o and that the mass of Jupiter is infinitely large compared to the mass of the probe. If a spacecraft approaches Jupiter with a velocity v_o and swings around Jupiter $180°$ as shown, what will its final velocity be?



79. In the late twentieth century the Pioneer 10 and 11 space probes were passing through the Kuiper belt, a region beyond Neptune that contains dust, ice and other small bodies, and which is thought to be the origin of some comets. Pioneer 10 had a mass of 240 kg, was travelling at a velocity of 12.2 km/s and carried a dish antenna with radius $R = 1.73$ m. The spacecraft was observed to decelerate with $a = 8 \times 10^{-10}$ m/s^2. It is known that some gaseous nebulae have densities of about 10^{-16} kg/m^3. Was the Pioneer deceleration consistent with this number?

4. ROTATION AND ANGULAR MOMENTUM REVIEW

This chapter is at the AP level. However, the definitions of angular velocity and acceleration in Section 2 are needed for subject-level problems in circular motion (Chapter 5).

1. Essential Terms

Mechanical systems are not always composed of masses that may be treated as point particles traveling along straight lines. Among the most important systems in physics are extended bodies that rotate. Such systems exhibit **angular momentum**, the rotational analog of linear momentum. Rotating bodies are characterized by the rotational analog of mass, **moment of inertia**. To change the angular momentum of a body requires the imposition of an external **torque**. Before discussing the concepts, however, it is essential to first define the variables that are appropriate to describing them, in particular, **angular velocity** and **angular acceleration**. The subject of rotation is more difficult than linear motion because the vectors involved are related in ways that are not initially obvious and require the cross product for their proper description. The SAT subject test assumes only the material in Section 2. For the AP Physics 1 and 2 exams everything in this chapter is fair game, although in practice very few questions about angular momentum appear on College Board exams.

2. Angular Velocity and Acceleration

Consider a circle of radius r (Figure 4.1a). The angle θ in radians subtending an arc length s is $\theta \equiv s/r$. When $s = r$, $\theta = 1$ radian. When $s = 2\pi r$, $\theta = 2\pi$. Thus, there are 2π radians in a full circle. Note that since a radian is defined as a length/length, it is *dimensionless*. Because there are 2π radians per $360°$, 1 radian $\approx 57° \approx 60°$.

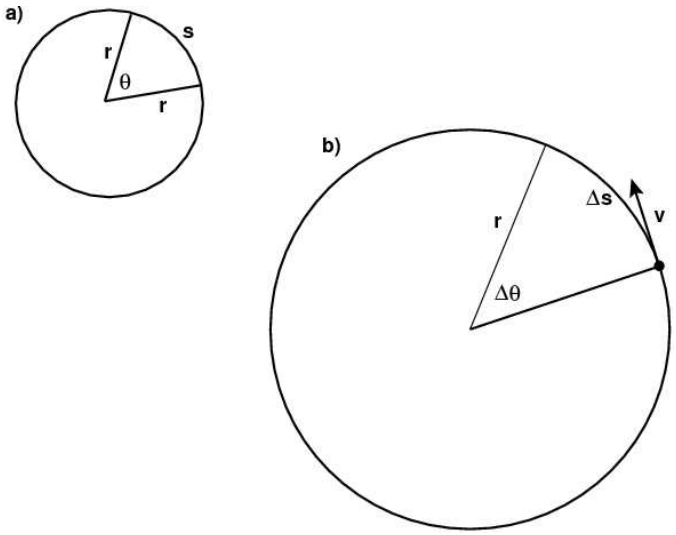


Figure 4.1

Suppose that a particle is traveling around the circle (Figure 4.1b). The vector $\mathbf{v}$ is termed the particle's **tangential velocity**; it is the linear velocity directed tangentially along the circumference. The particle moving at speed $v = |\mathbf{v}|$ will traverse a small linear distance Δs around the circle in a time Δt. Thus, $v = \Delta s/\Delta t$ as usual and is measured in $\mathrm{m/s}$. During the same period of time Δt the particle will also traverse the small angle $\Delta\theta$. The quantity $\Delta\theta/\Delta t$ is termed the **angular velocity** of the particle and is usually given the symbol ω ("omega"). *For a constant angular velocity*, we have

$$\theta = \omega t \quad (4.1)$$

which is the angular analogy of $s = vt$. From above, however, $\Delta\theta \equiv \Delta s/r$. Dividing by Δt gives

$$\omega \equiv \frac{\Delta\theta}{\Delta t} = \frac{1}{r}\frac{\Delta s}{\Delta t} = \frac{1}{r}v \quad .$$

Thus,

$$\omega = \frac{v}{r} \;\; ; \;\; v = \omega r \quad (4.2)$$

Although angular velocity really is a vector (Section 6), in these equations we have written only its magnitude, so we should call it angular speed, although physicists themselves rarely use that term. **Eq. (4.2) gives the relationship between tangential and angular speeds and is used all the time.**

Note that ω measures *angular displacement per unit time* and is measured in radians/s. However, because radians are themselves dimensionless, the dimensions of ω are $1/\mathrm{T}$ and its units are $1/\mathrm{s}$.

Also note that when an old-fashioned record completes one revolution it has gone through an angle $\theta = 2\pi$ radians. Two bugs sitting at r and $2r$ on the record complete one revolution in the same amount of time t and thus have the same angular velocity $\omega = 2\pi/t$. But according to Eq. (4.2) bug 2 has twice the linear velocity v as bug 1 and traverses twice the distance s. The time it takes a particle or bug to complete one circular orbit is termed the **period**, and is usually written T or P. We thus have $\omega = 2\pi/P$.

Linear acceleration is defined as $a = \Delta v/\Delta t$. Similarly, the magnitude of the angular acceleration α ("alpha") is

$$\alpha \equiv \frac{\Delta\omega}{\Delta t}.$$

Since $\omega = v/r$, we have

$$\alpha = \frac{a}{r} \quad ; \quad a = \alpha r \quad (4.3)$$

which gives the relationship between linear and angular accelerations. Note that angular acceleration has dimensions $1/\mathrm{T}^2$ and is measured in $1/\mathrm{s}^2$. Once again, bear in mind that in this section we have been talking about *scalars*: angular speeds and the magnitude of angular accelerations. Nevertheless, ω is usually referred to as the angular velocity in the same way that linear speed v is often referred to as velocity, even when we are not worried about the direction.

3. Moment of Inertia

In describing an extended, rotating system, the difficult problem is to devise an expression for the kinetic energy. Figure 4.2 illustrates the dilemma.

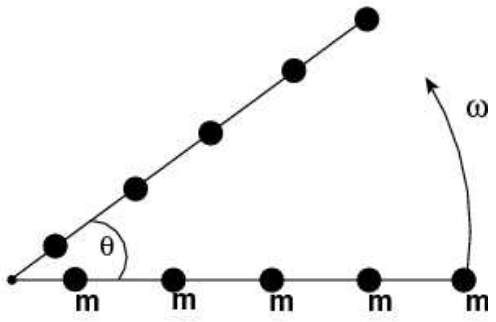


Figure 4.2

Imagine a rod, idealized as a series of masses m_i connected by thinner rods of negligible mass. The system is rotating counterclockwise with a constant angular speed ω about an axis on the left end. What is the kinetic energy of the rod? The total kinetic energy should be

$$KE_{tot} = \frac{1}{2}\sum_{i=1}^{N} m_i v_i^2 = \frac{1}{2}(m_1 v_1^2 + m_2 v_2^2 + \cdots + m_N v_N^2)$$

where v_i is the speed of each m_i and N is the total number of masses. Unfortunately, since each mass is at a different distance r_i from the rotation axis, each mass has a different v_i, and we don't immediately know what they are. However, because ω is constant for all the masses, we can write from Eq. (4.2) $v_i = \omega r_i$. Then the total KE becomes

$$KE_{tot} = \frac{1}{2}\sum_{i=1}^{N} m_i r_i^2 \omega^2$$

The quantity

$$I \equiv \sum_{i=1}^{N} m_i r_i^2 \quad (4.4)$$

is termed the **moment of inertia** of the system.[1] We can then write

$$KE_{tot} = \frac{1}{2} I\omega^2 \quad (4.5)$$

Eq. (4.5) gives the kinetic energy of an extended rotating object. **The moment of inertia *I* plays the role of mass for a rotating system**. However, the r_i's represent the distance of the various mass elements from a particular axis of rotation. To calculate I, therefore, one needs to know the mass distribution of mass about the given axis. The simplest system is a dumbbell, consisting of two equal masses m joined by a rod of negligible mass and length L (Figure 4.3a).

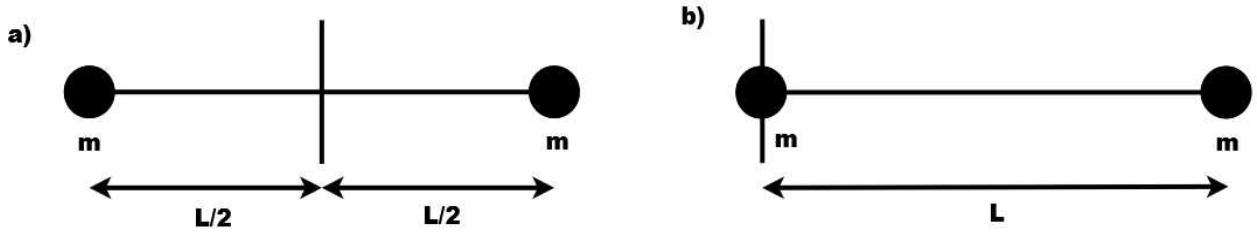


Figure 4.3

In the case where the masses are rotating around an axis through the center of the rod, the moment of inertia from Eq. (4.4) is $I = (mL^2)/4 + (mL^2)/4 = (mL^2)/2$. On the other hand, the same system rotating around an axis through one of the masses has moment of inertia $I = m0^2 + mL^2 = mL^2$, which is larger. For most systems, the calculation of the moment of inertia requires calculus and can often be rather difficult. We list a few common moments of inertia in the table at the end of this section.

The important thing to remember about moment of inertia is that *the dimensions of I are always* $[I] = ML^2$. Along with the simple example, this already tells you much: **Objects with their mass concentrated far from the axis of rotation have large moments of inertia; objects with their mass concentrated near the axis of rotation have small moments of inertia. Moments of inertia can always be written $I = CmR^2$, where C is a number ≤ 1. The maximum value 1 is attained when all the mass is concentrated at the edge of the object.**

[1] For continuous systems, the moment of inertia becomes $I = \int r^2 dm$.

When employing conservation of energy for rotating bodies one must include rotational energy:

$$KE = \frac{1}{2}mv^2 + \frac{1}{2}I\omega^2$$

where v is the linear velocity of the center of mass and ω is the rotational velocity around the center of mass

For a block or a ball, the center of mass will merely be the center of the object, as discussed in Section 3.13.

Example 4.1: Sphere and cylinder race. A solid sphere and a solid cylinder of the same material and same diameter roll down an inclined plane of height h without slipping. If they are started off at the same time at the top of the incline with zero velocity, which reaches the ground first?

The sphere must reach the ground first. Sitting at the top of the incline they each have potential energy $U = mg(h + R)$, where R is the radius of the two objects. As they roll down, potential energy is converted into kinetic energy, but from the below table, one sees that the cylinder has a larger moment of inertia and therefore more of the energy goes into rotation than into translational motion; hence it loses the race.

One can show this rigorously: When the sphere and cylinder reach the ground, they have potential energy $U = mgR$ (remember their centers of mass are at a distance R above the ground). The difference in potential from the top of the ramp is then mgh. By energy conservation:

$$mgh = \frac{1}{2}mv^2 + \frac{1}{2}I\omega^2.$$

The velocity v is the linear velocity of the object's *center*. If the object is not slipping on the ramp, then one can show $v = \omega R$. (**Exercise**: Demonstrate the previous statement.) In that case

$$mgh = \frac{1}{2}mv^2 + \frac{1}{2}I\frac{v^2}{R^2}.$$

Solving for v^2 gives

$$v^2 = \frac{2mgh}{m + \frac{I}{R^2}}.$$

Consequently, the larger the moment of inertia I, the smaller v. Although we have found v^2 at the instant the objects touch the ground, it is true for any distance h below the top of the ramp. The ratio of the speeds of the sphere to the cylinder at any h is

$$\frac{v_s}{v_c} = \sqrt{\frac{1 + \frac{1}{2}}{1 + \frac{2}{5}}} = \sqrt{\frac{15}{14}} > 1.$$

and so the sphere does win, barely.

Common moments of inertia, $I = CMr^2$ *for total mass M*

System	C
Dumbbell rotating around center	1/4
Dumbbell rotating around one end	1/2
Hoop rotating around center	1
Disk or solid cylinder rotating around center	1/2
Sphere rotating around any diameter	2/5
Thin rod rotating about axis through center, perpendicular to length	1/12
Thin rod rotating about axis through one end, perpendicular to length	1/3

4. Angular Momentum

Angular momentum is the rotational analog of linear momentum $\mathbf{p} = m\mathbf{v}$, and in analogy to linear momentum it measures the "oomph," or inertia, of a rotating system. Like linear momentum, angular momentum in an isolated system is conserved:

Conservation of Angular Momentum:

The total angular momentum of an isolated system never changes.

Unfortunately, rotational motion is much more complicated than linear motion; consequently, the definition of angular momentum is more complicated as well. Luckily, elementary courses and the AP test restrict themselves to the simplest situations, as we will.

The basic system exhibiting angular momentum is that of a particle of mass m moving at a constant speed v in a circle of radius r about an origin, as in Figure 4.1b. A ball on a string being whirled around your head is an example. For obscure reasons angular momentum, following the tradition of linear momentum, is represented by the letter **L**. The boldface indicates that, like linear momentum, angular momentum is a *vector*. In the case of the mass m performing circular motion, the *magnitude* of the angular momentum is simply:

$$L = mvr$$

$$L = m\omega r^2 \qquad (4.6)$$

Here we used Eq. (4.2) to write L both in terms of linear and angular velocity. As with linear momentum, no one's name is attached to angular momentum. Its dimensions are $[\mathrm{L}] = \mathrm{ML}^2/\mathrm{T}$ and it is measured in $\mathrm{kg\,m}^2/\mathrm{s}$.

Because L is conserved, in the system of a ball whirling around the circle, if the length of the string is shortened, the speed of the ball must increase in order to keep L constant. The classic example of angular momentum conservation is that of a figure skater on ice: pulling in her hands essentially decreases r and so her angular velocity must increase.

To discuss angular momentum for more complex systems one must enlist the moment of inertia, discussed in Section 3. From the second equation in Eq. (4.6) we see that dimensionally L is just $I\omega$. In fact, consider a simple system spinning around an "axis of symmetry," meaning an axis around which the mass is equally distributed—for instance a bicycle wheel spinning about its axle. In such systems the magnitude of the angular momentum is given exactly by

$$L = I\omega \quad (4.7)$$

where I is the moment of inertia about that axis. Since a bicycle wheel is a hoop with all its mass concentrated at the rim, its moment of inertia is $I = mr^2$and its angular momentum is $L = mr^2\omega$. The angular momentum of a spinning disk with $I = 1/2\ mr^2$ is therefore $L = 1/2\ mr^2\omega$.

5. Torque

In section 3.3 we defined force as the change in momentum per unit time. In complete analogy, **torque**, usually given the symbol τ ("tau"), is defined as the change in angular momentum per unit time:

$$\tau \equiv \frac{\Delta L}{\Delta t} \quad (4.8)$$

If an object is accelerating, a force is acting upon it. **If an object experiences an angular acceleration, a torque is acting upon it.** Just as a force is a push or a pull, a **torque is a twist**. If the moment of inertia is constant, Eqs. 4.7 and 4.8 immediately imply that $\tau \equiv \Delta L/\Delta t = I\ \Delta\omega/\Delta t$, or

$$\tau = I\alpha \quad (4.9)$$

where α is the angular acceleration. Just as in the equation $F = \Delta p/\Delta t$ where F refers to the total external force acting on an object, in Eq. (4.8) τ *refers to the total external torque acting on an object*. A torque may twist an object in the clockwise or counterclockwise direction. Therefore, in Eq. (4.9) $\tau = \tau_{cw} - \tau_{cc}$, the difference between any clockwise or counterclockwise torques present.

Notice that the dimensions of torque are $\mathrm{ML^2/T^2}$, which are the same as the dimensions of work or energy. This can cause confusion, because although the dimensions are the same, *torque and energy are not the same*. It does help, however, in remembering that the units for torque are kg $\mathrm{m^2/s^2}$, or newton-meters. It also helps motivate the alternative definition of torque.

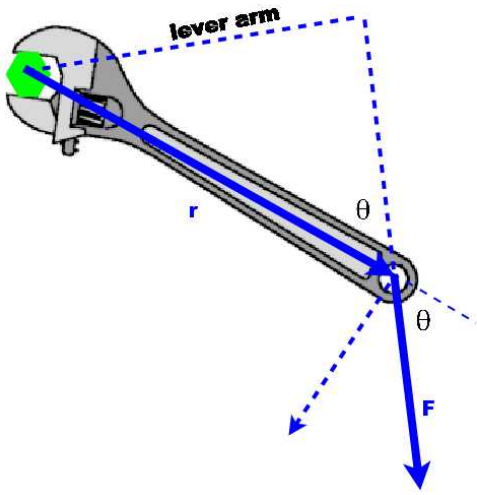


Figure 4.4

If one wants to tighten a bolt, as in Figure 4.4, one exerts a force F at the end of a wrench. From daily life you know that the longest possible wrench gives the greatest tightening experience. This is called torque. (Torque wrenches are not called torque wrenches for nothing.) However, the amount of torque you can exert on the bolt depends not only on the wrench's length r, but also on the angle θ at which the force is exerted. If $\theta = 0$, the only thing you have done is pull the wrench off the bolt. Clearly, you maximize the torque when $\theta = 90°$. The magnitude of the torque is thus sensibly defined as

$$\tau = Fr \sin\theta \qquad (4.10)$$

where θ is the *smallest* angle between the vectors **r** and **F**. From the diagram, you see that $r \sin\theta$ is the perpendicular distance from the axis of rotation (the bolt) to the point of application of **F**. This perpendicular distance is termed the **lever arm**. Maximizing the torque requires maximizing the lever arm, which occurs at $\theta = 90°$. Then **F** is perpendicular to **r** and the length of the lever arm is r—the full length of the wrench. However, the figure also shows that the component of **F** perpendicular to the wrench has magnitude $F \sin\theta$, which acts at the full distance r from the bolt. Therefore, you can think of torque either as

$$\text{torque} = \text{force} \times (\text{lever arm})$$

or

$$\text{torque} = (\text{perpendicular component of force}) \times (\text{distance to pivot}).$$

The two are equivalent. Furthermore, and although it may not be obvious, we show in Section 6 that the two definitions of torque, Eq. (4.8) and Eq. (4.10) are also equivalent.

Example 4.2: Atwood machine. Consider two masses hanging from a frictionless pulley with radius R and mass M, as shown. Such a configuration is known as an Atwood machine. What is the acceleration of the masses and the tensions in the string?

Advice: Do Problem 3.22 first.

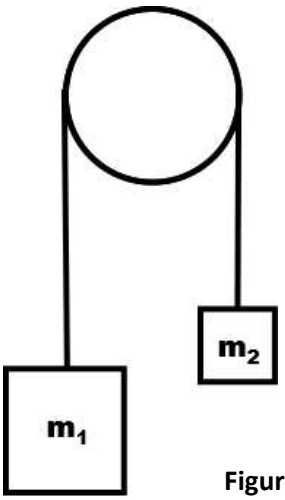


Figure 4.5

The first thing to realize is: **When the pulley in such a system has nonzero mass the tension in the string is no longer constant.** It is easy to see why. If the pulley has mass M it also has a nonzero moment of inertia $I = 1/2\ MR^2$. As the masses accelerate, the pulley undergoes an angular acceleration α, and so by Eq. (4.9) the torque τ must be nonzero. Because the string is the only thing producing the torque and acts at right angles to the axle, Eq. (4.10) tells us that the total torque is $\tau = (T_L - T_R)R$, where the radius R of the pulley is the lever arm and T_L and T_R are the tensions on the left and right sides of the pulley, respectively. If T_L and T_R were the same, the torque would be zero; since the torque cannot be zero, the tensions must be unequal.

In solving this problem, it may help to visualize the masses stretched out in a line, with gravity pulling m_1 to the left and m_2 to the right. T_R then acts to the left and T_L acts to the right. Call to the left positive. Newton II for m_1 (the 8 kg mass) becomes:

$$m_1 a_1 = m_1 g - T_L$$

Newton II for m_2 is:

$$m_2 a_2 = -m_2 g + T_R$$

However, the string provides a **physical constraint** that requires $a_2 = a_1$. Call the joint acceleration a. Then adding the two equations gives

$$(m_1 + m_2)a = (m_1 - m_2)g + T_R - T_L$$

We know from above that $\tau = (T_L - T_R)R = I\alpha = 1/2\ MR^2\alpha$. If the string does not slip on the pulley we can say that $\alpha = a/R$ (see Example 4.1). Then $T_L - T_R = 1/2\ Ma$ and

$$a = \frac{(m_1 - m_2)g}{m_1 + m_2 + \frac{1}{2}M}$$

Does this answer pass the reality check? Yes. If $m_1 = m_2$, then $a = 0$, as expected, since the force of gravity is acting on the two masses equally. Further, a is less when $M \neq 0$ than when $M = 0$; with a nonzero I, the pulley has inertia and adds an effective mass to the system, lessening any acceleration. One can now find the tensions by merely plugging the expression for a back into the top two equations.

Example 4.2 is not unrelated to the important concept of *gravitational torque*.

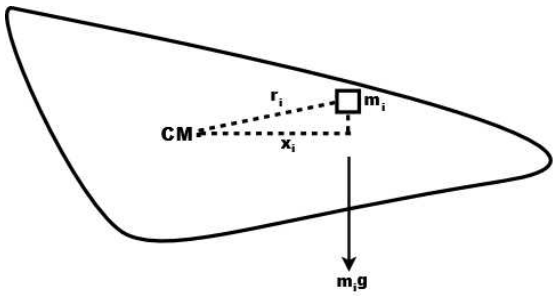


Figure 4.6

Consider the irregular object shown in Figure 4.6. A given mass element m_i lies at a distance r_i from the center of mass and has a lever arm x_i. Gravity acts downward on the mass element with a force $m_i g$. The torque on this mass element is therefore $\tau_i = m_i g x_i$. The total torque on the entire object will be

$$\tau = Mg \sum_i \frac{m_i x_i}{M}$$

where we have multiplied top and bottom by the object's total mass M. However, from Eq. (3.18), we see that the summed quantity is, by definition x_{CM}, the distance of the center of mass of the object from the chosen origin. Therefore, the total gravitational torque on any object is merely

$$\tau = Mgx_{CM} \quad (4.11)$$

This equation says that, in terms of torque, one can regard any object as a point mass; gravitational torque acts on this object through the center of mass, whose lever arm is the distance x_{CM} from the point chosen as zero. If the CM itself is chosen as the origin, then the total torque is zero and the object is balanced.

The existence of torques forces us to modify the definition of **equilibrium** mentioned in Section 3.3.

A system is in equilibrium if the sum of all the forces and all the torques acting on it is zero.

By this definition, a system rotating at a constant angular velocity ω and whose center of mass is moving at constant linear velocity v is in equilibrium. If we require that it not be moving at all, the system is said to be in **static equilibrium**.

Example 4.3: Ladder against wall. This problem incorporates many of the concepts discussed in the present chapter.

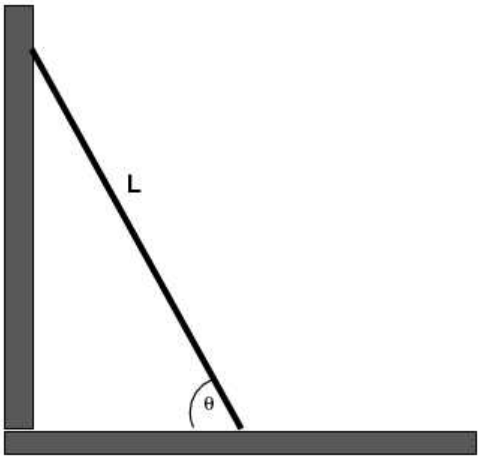


Figure 4.7

Consider a ladder of length L leaning against a frictionless wall at an angle θ, as shown in Figure 4.7. What is the minimum value of the coefficient of static friction μ between the ladder and the floor required to keep it from slipping?

The first thing to realize in solving this problem is that for the ladder to be in static equilibrium, *the sum of all the forces and torques acting on it must be zero*. Call the mass of the ladder m and refer to Figure 4.8 below.

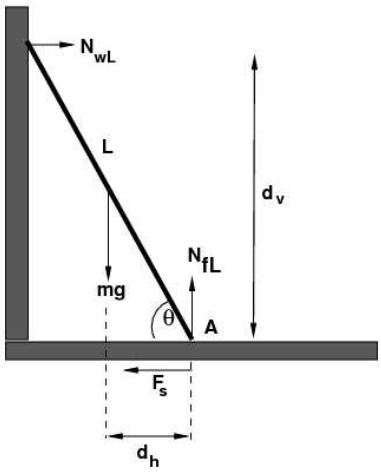


Figure 4.8

There are only two x-forces on the ladder: the normal force exerted by the wall on the ladder, labeled N_{wL}, which acts to the right, and the force of static friction, F_s, which acts to the left to keep the ladder from slipping. We can set $F_s = \mu mg$. There are only two y-forces on the ladder: the force of gravity mg acting down and the normal force N_{fL} that the floor exerts upward. Setting the x- and y-forces equal to zero gives

$$N_{wL} - \mu mg = 0$$
$$mg - N_{fL} = 0$$

To determine the two normal forces, we use the condition that the total torque on the ladder must equal zero. **To calculate the torques, pick any convenient axis of rotation.** (This is usually where something is zero.) Here we choose the point at which the ladder touches the floor, A. From the figure, the counterclockwise torque is mgd_h, where the horizontal distance d_h is the lever arm from point A. The clockwise torque is $N_{wL}d_v$, where d_v is the lever arm through which the normal force N_{wL} acts. We must therefore have $mgd_h = N_{wL}d_v$. In terms of θ this gives

$$\frac{mgL}{2}\cos\theta = N_{wL}L\sin\theta$$

which tells us that $N_{wL} = mg\cot\theta/2$. Plugging this expression into the top equation gives

$$\mu = \cot\theta/2$$

If $\theta = 60°$, for example, $\mu = 0.29$.

To solve this problem, we had to do it by the book: set all components of the forces equal to zero, as well as the torques. It was also important to realize that a normal force does not necessarily act up; N_{wL}acted to the *right*. Finally, it is worth bearing in mind that μ is dimensionless and in all problems of this sort it ends up being the tan or cot of some angle.

6. Vector Nature of Angular Momentum

In sections 2-5 we treated angular velocity, angular acceleration, angular momentum and torque as scalars, but in fact all are vector quantities. Unfortunately, they do not point in intuitive directions and are related to each other in complicated ways, which students often have difficulty visualizing. The tool for describing the relationships among all these quantities is the cross product from Chapter 1.

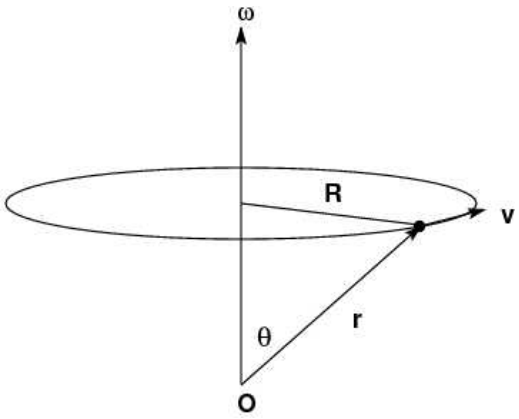


Figure 4.9

Figure 4.9 shows a particle moving with tangential velocity **v** in a circle at a distance R from an axis. The vector **r** is the particle's displacement vector from the origin O.

The direction of the angular velocity $\boldsymbol{\omega}$ is defined as the direction of advance of a right-hand screw turned in the same sense as the velocity of the particle. This is sometimes called **right-hand rule No. 2**. In other words, **curl the fingers of your right hand around in the direction the particle is moving. The direction your thumb points is the direction of $\boldsymbol{\omega}$.**

Note that $R = r\sin\theta$. From Eq. (4.2), therefore, $v = r\omega\sin\theta$. From Eq. (1.3) we recognize this as the magnitude of the cross product between $\boldsymbol{\omega}$ and **r**. As a vector equation

$$\mathbf{v} \equiv \boldsymbol{\omega} \times \mathbf{r} \quad (4.12)$$

Because the cross product is noncommutative, you must get the factors in the correct order. (**Exercise:** check that Eq. (4.12) correctly describes Figure 4.9.)

We also had from Eq. (4.6) that the angular momentum for a particle was $L = mvr$, or $L = pr$. As a vector equation,

$$\mathbf{L} \equiv \mathbf{r} \times \mathbf{p} \quad (4.13)$$

From Eq. (4.10) we had $\tau = Fr\sin\theta$. As a vector equation this becomes

$$\boldsymbol{\tau} \equiv \mathbf{r} \times \mathbf{F} \quad (4.14).$$

But in Eq. (4.8) we also defined torque as the change in angular momentum per unit time: $\boldsymbol{\tau} \equiv \Delta\mathbf{L}/\Delta t$. Therefore, for the two definitions to be consistent we must have $\Delta\mathbf{L}/\Delta t = \mathbf{r} \times \mathbf{F}$. For those who know calculus, the demonstration can be found in the footnote.[2]

Exercise: Using the right-hand rule, form the cross product in Eq. (4.13) for the system in Figure 4.9.

Notice that **L** does *not* point in the direction of $\boldsymbol{\omega}$. However, in Eq. (4.7) we said that $L = I\omega$. This suggests that $\mathbf{L} = I\boldsymbol{\omega}$ is *not* in general the correct vector relationship between **L** and $\boldsymbol{\omega}$. Indeed, this is exactly why we stated that $L = I\omega$ is true only around an axis of symmetry. Add another mass to Figure 4.9 on the opposite side of the circle from the one drawn and convince yourself that when that is done $L = I\omega$.

Further aspects of the vector nature of angular momentum will be explored in the problems.

---

[2] The two definitions of torque require that $\mathbf{r} \times \mathbf{F} = d(\mathbf{r} \times \mathbf{p})/dt$. The cross product obeys the product rule for differentiation as long as the order of the factors is preserved. Thus, $d(\mathbf{r} \times \mathbf{p})/dt = d\mathbf{r}/dt \times \mathbf{p} + \mathbf{r} \times d\boldsymbol{p}/dt$. However, $d\mathbf{r}/dt \equiv \mathbf{v}$ and since $\mathbf{p} = m\mathbf{v}$, $\boldsymbol{p}$ and **v** are in the same direction. From Chapter 1, we know that the cross product of two vectors in the same direction is zero, and so $d(\mathbf{r} \times \mathbf{p})/dt = \mathbf{r} \times d\boldsymbol{p}/dt$. But by definition $d\boldsymbol{p}/dt = \mathbf{F}$. Therefore, $d(\mathbf{r} \times \mathbf{p})/dt = \mathbf{r} \times \mathbf{F}$, as desired.

ROTATION AND ANGULAR MOMENTUM PROBLEMS

The problems in this chapter are all at the AP level.

1. Which of the following could be a unit of angular momentum?

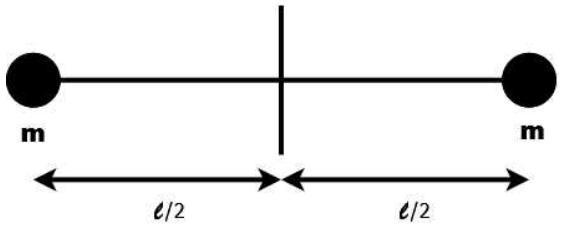
   i. N m s ii. $\sqrt{\text{J kg}}$ m iii. W s^2 iv. J/m v. W s^2/m

   (A) i
   (B) ii
   (C) iii
   (D) i, ii and iii
   (E) i, ii, iii, iv, v

2. The *Enterprise* is in a circular orbit around the Sun when the crew discovers a nuclear bomb aboard. As in all science fiction films, the crew decides to jettison the bomb and let it fall into the Sun, thus saving the ship. Which statement is true?

   (A) The bomb will fall into the Sun as planned, saving the *Enterprise*.
   (B) The bomb will follow the *Enterprise* in orbit until it explodes.
   (C) The bomb will gradually fall behind the *Enterprise* and explode at a safe distance.
   (D) The crew could move the bomb to a lower orbit by attaching a rocket to it.
   (E) (B) and (D)

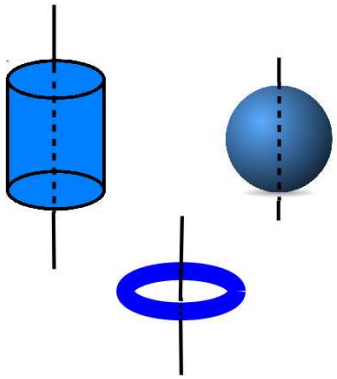
3. A dumbbell consisting of two identical masses m located at a distance $\ell/2$ from an axis of rotation, as shown below, is rotating at an angular velocity ω. The mass of the rod is negligible.



   What is the dumbbell's angular momentum?

   (A) $m\ell^2\omega/2$
   (B) $m\ell\omega/2$
   (C) $m\ell^2\omega$
   (D) $m\ell^2\omega^2/2$
   (E) $m\ell^2\omega/4$

4. A hollow sphere, a solid sphere, a hoop and a solid cylinder all have the same mass M and radius R.

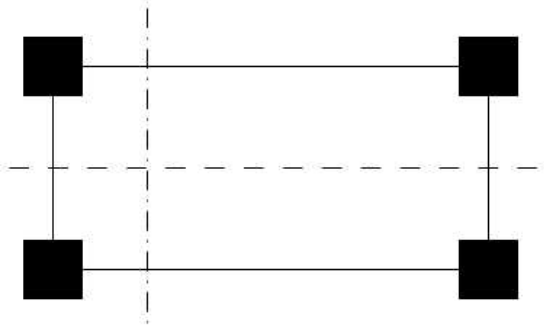


Rank the following moments of inertia from highest to lowest.

i. The hollow sphere spinning around a diameter
ii. The solid sphere spinning around a diameter
iii. The hoop spinning around an axis through its center
iv. The solid cylinder spinning around an axis through its center

(A) ii, iv, i, iii
(B) iv, ii, i, iii
(C) iii, iv, ii, i
(D) i, iii, ii, iv
(E) iii, i, iv, ii

Questions 5 - 6 refer to the following information:



Four blocks are connected by thin rods of negligible mass. All four blocks have the same mass m. The length of the horizontal rods is L. The length of the vertical rods is ℓ. Ignore the size of the blocks themselves.

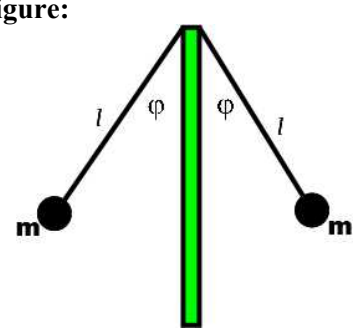
5. If the horizontal axis passes midway between the blocks, as shown, then the moment of inertia around this axis is equal to

(A) $2m\ell^2$
(B) $m\ell^2$
(C) $m\ell^2/2$
(D) $m\ell^2/4$
(E) $m\ell/6$

6. If the vertical axis is at a distance $L/4$ from the left end, then then the moment of inertia around this axis is equal to

(A) $mL^2/3$
(B) $mL^2/4$
(C) mL^2
(D) $5mL^2/4$
(E) $5mL^2/6$

Questions 7 - 8 refer to the following figure:



7. Two identical balls of mass m are rotating around an axle on thin rods at an angular velocity ω, as shown. The length of the rods, which are of negligible mass, is ℓ and they make an angle φ with the vertical. The angular momentum of the system is

   (A) $m\ell^2\omega/2$
   (B) $2m(\ell \sin\varphi)^2\,\omega$
   (C) $2m(\ell \cos\varphi)^2\,\omega$
   (D) $2m\ell^2\omega$
   (E) $2\,m^2\ell\omega$

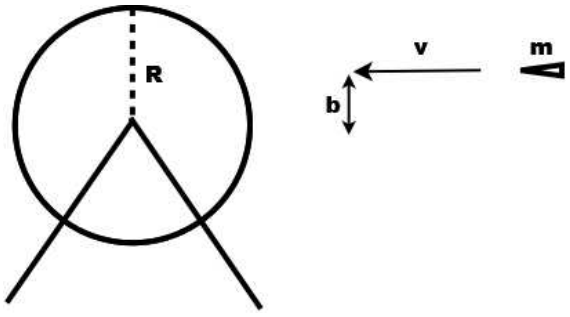
8. Consider just one of the arms. What is the torque exerted by gravity on the arm?

   (A) $mg\ell$
   (B) $mg(\ell \sin\varphi)^2$
   (C) $mg(\ell \cos\varphi)^2$
   (D) $mg\ell \sin\varphi$
   (E) $mg\ell \cos\varphi$

9. A simple pendulum is composed of a mass m attached to a massless string of length ℓ. The other end of the string is attached to the ceiling. A physical pendulum consists of a thin rod of uniform density with mass m and length ℓ. One end is attached to the ceiling by a frictionless bearing. Both the simple pendulum and the physical pendulum are displaced by an angle θ from the vertical and released from rest. Which pendulum reaches the equilibrium ($\theta = 0$) position first?

   (A) The simple pendulum
   (B) The physical pendulum
   (C) They both reach the equilibrium position at the same time.
   (D) The answer depends on the mass of the pendulums.
   (E) None of the above

10. A bullet of mass m is fired with a velocity v into a wooden disk of mass M and radius R that is free to rotate, as shown.



    The bullet lodges itself in the disk at a height b above the wheel's center and the system begins to rotate with an angular velocity ω. Assume $M \gg m$. In order to predict ω, one needs

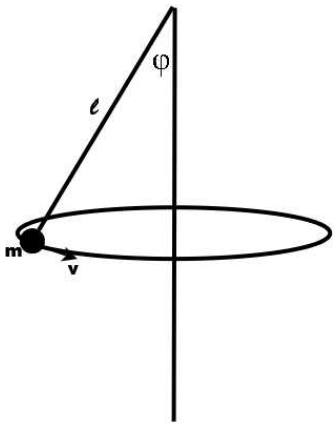
    (A) conservation of energy.
    (B) conservation of momentum.
    (C) conservation of angular momentum.
    (D) conservation of energy and conservation of momentum.
    (E) conservation of energy and conservation of angular momentum.

11. A window washer of $M = 75$ kg is taking a lunch break and sitting 4/3 m from the left edge of a 5- meter-long board that hangs from two cables. The mass m of the board is 20 kg. If T_1 is the tension in the right cable and T_2 is the tension in the left cable, what are their values?

    (A) $T_1 = 200$ N; $T_2 = 650$ N
    (B) $T_1 = 300$ N; $T_2 = 650$ N
    (C) $T_1 = 650$ N; $T_2 = 300$ N
    (D) $T_1 = 650$ N; $T_2 = 200$ N
    (E) $T_1 = 450$ N; $T_2 = 400$ N

Questions 12 - 16 refer to the following information:

A ball of mass m is attached to the ceiling by a string of length ℓ, which makes an angle φ with the vertical as the ball whirls around in a circle with constant tangential speed v.



12. The angular velocity vector $\boldsymbol{\omega}$ points

    (A) in the direction of **v**.
    (B) opposite the direction of **v**.
    (C) down.
    (D) up.
    (E) radially outward.

13. At the instant shown in the figure, the angular momentum vector **L** points

    (A) in the direction of **v**.
    (B) toward the point of attachment, along the string.
    (C) outward from m at an angle φ above the horizontal.
    (D) in the direction of $\boldsymbol{\omega}$.
    (E) outward from m at an angle φ from the vertical.

14. In general, the angular momentum vector **L**

    (A) changes direction and so there must be a torque on the ball.
    (B) has constant magnitude and so there cannot be a torque on the ball.
    (C) changes in magnitude and so there is a torque on the ball.
    (D) changes in magnitude but stays constant in direction, so there is no torque on the ball.
    (E) changes in both magnitude and direction but in such a way as there is no torque on the ball.

15. The magnitude of the ball's angular momentum is

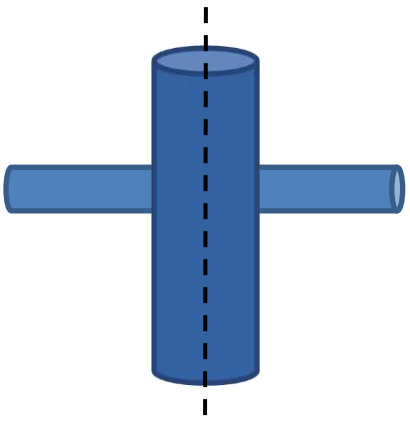
    (A) $mv\ell$
    (B) $m(\ell \sin\varphi)^2\,\omega$
    (C) $m(\ell \cos\varphi)^2\,\omega$
    (D) $m\ell^2\omega$
    (E) $m^2\ell\omega$

16. The product $I\omega$ for the ball equals

(A) $m\ell^2\omega/2$
(B) $m(\ell \sin\varphi)^2\,\omega$
(C) $m(\ell \cos\varphi)^2\,\omega$
(D) $m\ell^2\omega$
(E) $m^2\ell\omega$

Free response questions

17. A figure skater can be idealized as composed of two solid cylinders of uniform density crossed at right angles to one another.



Assume that the skater is spinning with angular speed ω around a vertical axis passing through the center of her body. The mass of the large cylinder is M; it has radius R and length L. The mass of the small cylinder is m; its radius is r and its length is ℓ.

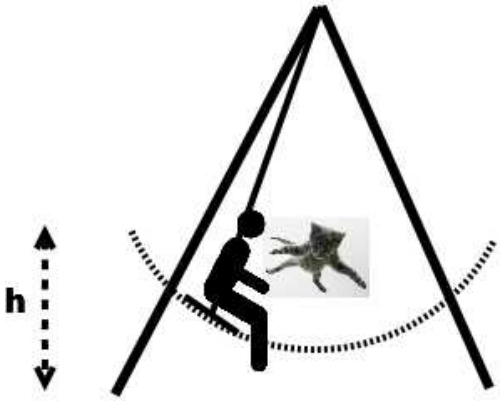
a) What is the moment of inertia of the large vertical cylinder in terms of the given quantities?

b) What is the moment of inertia of the horizontal cylinder in terms of the given quantities?

c) What is the figure skater's initial angular momentum?

d) If $m = 0.1M$, $\ell = 8R$ and the skater's arms can be totally retracted to her body, what is her new angular speed in terms of ω?

18. You are swinging on your backyard swing from an initial height h, as shown in the figure below. Assume that the mass of the swing and rope are negligible compared to your own mass, and that the swing is frictionless.



a) Just as you reach the lowest point of your swing, your pet cat, Beta, who has been patiently sitting by the side of the swing, suddenly jumps onto your lap. You expect your maximum height on the far side of the swing to be

i. equal to h.
ii. less than h.
iii. greater than h.

Justify your answer qualitatively (no calculations required).

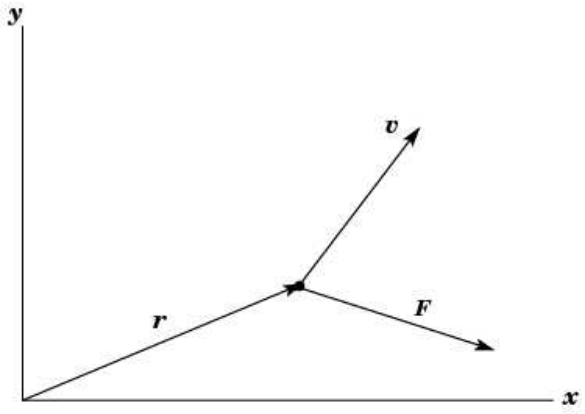
b) Explain how you would calculate the height that you expect to reach, call it h_1, without doing any actual calculations.

c) As you are swinging back toward your original position, just as you reach the lowest point of the swing, Beta jumps off. You expect your final height to be

i. equal to h_1.
ii. greater than h_1 but less than h.
iii. equal to h.

Justify your answer qualitatively (no calculations required).

19. A particle of mass m with a velocity **v** is at a position **r**, as shown below. A force **F** acts on the particle.



Let $m = 2$ kg, $\mathbf{v} = 5$ m/s at an angle $\alpha = 60°$ above the horizontal. Let $\mathbf{r} = 3$ m at an angle $\beta = 30°$ above the horizontal and let $\mathbf{F} = 4$ N at an angle $\gamma = 15°$ below the horizontal.

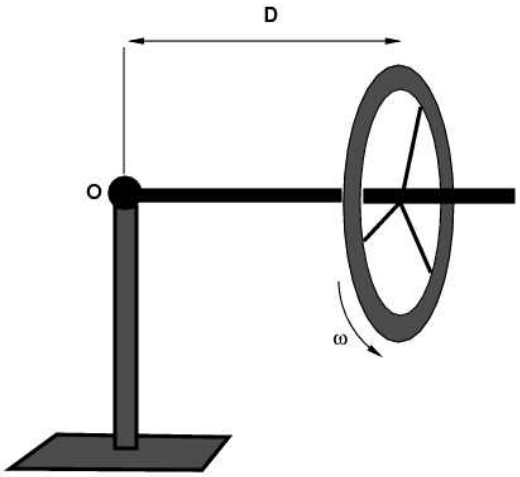
a) For a particle in arbitrary motion around the origin, find the components of the angular momentum in terms of v_x, v_y, v_z and x, y, z.

b) For the above particle, whose motion is confined to the xy-plane, show that there is only a z-component of the angular momentum.

c) Find the angular momentum around the origin of the particle.

d) Find the torque acting on the particle.

20. A gyroscope is essentially a spinning wheel attached to an axle that is free to rotate in any direction about a pivot. In the figure below, a gyro wheel of mass M and radius R spins with angular velocity ω in a counterclockwise direction when viewed from the right. The distance between the wheel and the pivot O is D. Assume that the axle is of negligible mass and that the wheel's entire mass is concentrated around the rim.



a) Is there an axis of symmetry in the problem? If so what is it?

b) What is the magnitude of the angular momentum of the gyroscope? Draw the direction the angular momentum vector is pointing on the diagram.

c) Are there any external forces acting on the gyro wheel? If so, draw them on the diagram.

d) What is the magnitude of any torque acting on the wheel? Draw the direction of the torque on the diagram.

e) Does the gyroscope move in any direction? If so which?

5. CIRCULAR MOTION, SIMPLE HARMONIC MOTION AND WAVES REVIEW

Section 7 is at the AP level.

1. Essential Terms

The topics covered in this chapter also fall under the general heading of mechanics. As its name surely implies, **circular motion** concerns the special case of objects traveling in a circle, which gives rise to the phenomenon of **centripetal force**. In Section 4.2, we introduced several quantities necessary for a discussion of circular motion, in particular **angular velocity** and **angular acceleration**. The same quantities also appear in any discussion of **simple harmonic motion**—the motion executed by springs and pendulums—and in any discussion of **waves**. All these phenomena are collectively said to undergo **periodic motion** and therefore we have grouped them together. Few questions on College Board exams directly concern circular motion. However, the basic concepts in Sections 2-3 are necessary to understand simple harmonic motion and waves. The SAT physics subject test and the AP physics exams have a substantial number of questions on these topics.

2. Kinematic Equations for Circular Motion

Uniform circular motion refers to an object traveling in a circle at a *constant speed* v. In that case the angular acceleration α is zero. In Section 4.2 we established the relationships between linear and angular variables for this case: $s = \theta r$; $\theta = \omega t$; $v = \omega r$.

However, there is no reason that the angular acceleration must be zero. In the general case $a = \alpha r$ and neither ω nor v will remain constant. Nevertheless, as long as α is constant, one can easily handle this situation. Merely replacing s, v and a by θr, ωr and αr respectively in the one-dimensional kinematical equations (2.4)-(2.6) allows us to write directly their analogs for circular motion:

$$\omega_f = \alpha t + \omega_o \quad (5.1)$$

$$\theta = \frac{1}{2}\alpha t^2 + \omega_o t \quad (5.2)$$

$$2\alpha\theta = \omega_f^2 - \omega_o^2 \quad (5.3)$$

Equations (5.1)-(5.3) should be regarded as scalar equations; they do not involve the angular vectors discussed in Section 4.6, but assume only the angular speed and magnitude of the acceleration. (As already stated, however, we usually refer to ω and α as angular velocity and angular acceleration.) In analogy to their linear counterparts, **Equations (5.1)-(5.3) hold only for constant α.**

Example 5.1: The Atwood machine (again). Consider once more the Atwood machine with two masses, m_1 and m_2, and a pulley of radius R and mass M. If the masses are released from rest, what will be the angular acceleration α of the pulley, its angular velocity ω and the angle through which it has rotated after a time t?

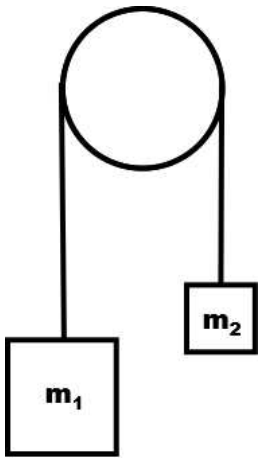


Figure 5.1

Because the pulley in this example has a nonzero mass, it must have a nonzero moment of inertia, which can be taken to be $I = 1/2\ MR^2$. In Example 4.2 we showed that the acceleration of the masses in such a device was

$$a = \frac{(m_1 - m_2)g}{m_1 + m_2 + \frac{1}{2}M}$$

Assuming the string is not slipping, the angular acceleration of the pulley is simply $\alpha = a/R$. We can then plug into Eqs. (5.1) and (5.2) to find the desired quantities. If $m_1 = 8$ kg; $m_2 = 2$ kg ; $M = 4$ kg ; $R = 20$ cm, then $a = 5\ \text{m/s}^2$ and $\alpha = 25\ \text{rad/s}^2$. After $t = 2$ s, Eq. (5.1) gives $\omega = 50\ \text{rad/s}$ and Eq. (5.2) gives $\theta = 50$ rad, or slightly more than 8 revolutions.

3. Period and Frequency

Several other quantities appear in any discussion of periodic motion. For an object in circular motion, the **period** T is merely the time necessary for the body to complete one full revolution. Because time is distance/velocity and the circumference of a circle is $2\pi r$,

$$T = \frac{2\pi r}{v} \quad (5.4)$$

for tangential velocity v. The inverse of the period is termed the **frequency** of rotation and is usually denoted by ν ("nu") or f:

$$f = \frac{1}{T} \quad (5.5)$$

Frequency is always the number of counts of something per unit time. *For uniform circular motion, frequency is the number of times per second an object completes an orbit.* If an object takes five seconds (the period) to complete an orbit, then the frequency is five orbits per second. Usually we speak of frequency in **cycles per second** or revolutions: "A vinyl LP spins at 33 1/3 revolutions per minute," which means that the frequency is about one-half revolution per second. Note that although we insert the word "cycles" or "revolutions," **the dimensions of frequency are 1/time**. 1 cps has been given the name hertz: 1 Hz $\equiv$ 1 cps.

We can also express the same quantities in terms of the angular velocity ω. Since $v = \omega r$, Eq. (5.4) becomes

$$T = \frac{2\pi}{\omega} \quad (5.6)$$

Plugging this expression for T into Eq. (5.5) shows that

$$\omega = 2\pi f \quad (5.7)$$

This relationship also follows by remembering that since there are 2π radians in a circle, $1\text{ cps} = 2\pi\text{ rad/s}$. When speaking of periodic motion, ω is referred to as **angular frequency**. In the case of uniform circular motion, it is the same as angular velocity.

Tip: All these expressions occur so frequently in physics that they should be memorized. Even better, know what they mean.

4. Centripetal Acceleration and Centripetal Force

Even if an object is moving around a circle at constant speed, its *velocity* is constantly changing direction. Because the velocity vector is changing, by definition the object *must* be accelerating. The acceleration of an object executing uniform circular motion is termed **centripetal acceleration**, from the Greek for "inwardly directed." Centripetal acceleration is correctly named: **The centripetal acceleration of an object is always directed toward the center of the circle.**

The behavior can be seen from Figure 5.2. An object is traveling in a circle with constant linear speed v. The tangential velocity vector initially points in the direction $\mathbf{v}_1$. A short time later, it points in the direction $\mathbf{v}_2$. We displace $\mathbf{v}_2$ parallel to itself until it is at the same starting position as $\mathbf{v}_1$; we can then compare the two vectors and draw the difference vector $\Delta\mathbf{v}$. As you see, $\Delta\mathbf{v}$ points in the general direction of the center of the circle. (**Exercise:** By drawing vectors, convince yourself that as the difference in time Δt gets very small such that $\mathbf{v}_1$ and $\mathbf{v}_2$ are pointed in almost the same direction, $\Delta\mathbf{v}$ points exactly toward the center of the circle.) Because the definition of acceleration is $\mathbf{a} = \Delta\mathbf{v}/\Delta\text{t}$, if $\Delta\mathbf{v}$ points toward the center of the circle, so does $\mathbf{a}$.

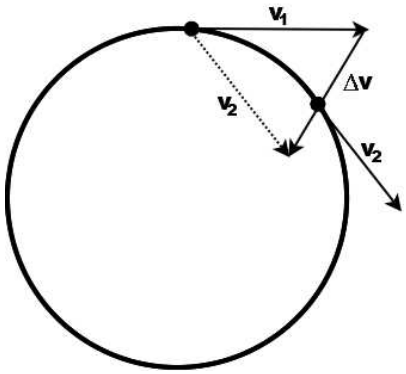


Figure 5.2

The magnitude of the centripetal acceleration can be guessed by dimensional analysis. Dimensionally, $a = v/t$. The only distance scale in a circle is its radius r and the only time we can make from the radius and velocity is $t = r/v$. Therefore, we should have $a = v^2/r$. This is in fact the exact answer. (**Exercise:** On Figure 5.2, draw in the central angle $\Delta\theta$ between the tail of v_1 and the tail of v_2. Show that for small angles $\Delta\theta \approx \Delta v/v$, where v is the constant length of the velocity vector. Use $\Delta\theta = \omega\Delta t$ and the relationship between ω and v to show that the centripetal acceleration is as stated.)

Of course, $F = ma$ still holds and so

Centripetal Acceleration

$$a = \frac{v^2}{r} = \omega^2 r \quad (5.8)$$

Centripetal Force

$$F = \frac{mv^2}{r} = m\omega^2 r \quad (5.9)$$

Note: Centripetal acceleration is *not* α, which is the change of angular speed ω with time. Centripetal acceleration exists even when ω is constant and therefore $\alpha = 0$.

Warning: Centripetal force is NOT a fundamental force of nature. Any force that causes an object to move in a circle becomes a centripetal force. If more than one force causes an object to move in a circle, then the centripetal force is the RESULTANT of those forces.

Example 5.2: Gravity as a centripetal force. Assume that the Moon revolves around the Earth at a constant speed in a circle of radius r. According to the above discussion, the Earth's gravity *must* produce a centripetal force on the Moon. From Newton's law of Gravity, Eq. (3.4), and Eq. (5.9) we then have

$$\frac{Gm_e m_m}{r^2} = \frac{m_m v^2}{r}$$

where m_e is the mass of the Earth and m_m is the mass of the Moon. The mass of the Moon appears on the right because it is the *Moon* that is accelerating. However, m_m cancels from both sides of the equation. Using Eq. (5.4) to eliminate v^2 in favor of T^2 gives

$$T^2 = \frac{(2\pi)^2}{Gm_e} r^3 \quad (5.10)$$

Because everything here except T and r is constant, Eq. (5.10) says that the square of the period of the Moon around the Earth (or a planet in orbit around the Sun) is directly proportional to the cube of the orbital radius. This is **Kepler's third law of planetary motion** for circular orbits, the foundation of modern physics.

Assume, furthermore, that the Moon is at a distance of $r = 3.8 \times 10^8$ m from the Earth and that it completes an orbit every 29.5 days. Using Eq. (5.4) to find the speed gives $v = 9.37 \times 10^2$ m/s, and using Eq. (5.8) to find the centripetal acceleration gives $a_c \approx 2.3 \times 10^{-3}$ m/s^2.

Isaac Newton did not know the gravitational constant G in the above equations, but he did believe that $g = Gm_e/R^2$, where R is the radius of the Earth, and he also believed that the acceleration of Earth's gravity at the distance to the Moon should be $a_g = Gm_e/r^2$. Dividing a_g by g gives

$$a_g = \frac{R^2}{r^2} g$$

which is merely the statement that the acceleration of gravity diminishes with the square of the distance. Taking the radius of the Earth to be $R = 6.4 \times 10^6$ m and the above value for r yields $a_g \approx 2.8 \times 10^{-3}$. (**Exercise:** Fill in the steps to these calculations.)

Note that the centripetal acceleration a_c is approximately equal to the Earth's gravitational acceleration at the distance to the Moon, a_g. It was this "agreement" that led Newton to believe that his law of gravity was correct.

Example 5.3: Tension as a centripetal force. A ball attached to the ceiling by a string revolves in a circle around an axis at an angle θ from the vertical, as shown. What is θ in terms of the radius of the circle r and the ball's constant tangential speed v?

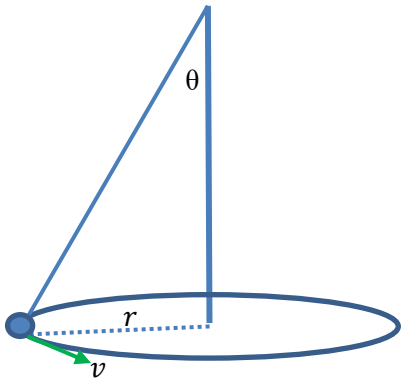


Figure 5.3

Because the string is holding up the ball and forcing it to move in a circle, there must be a tension T in the string. (Draw it.) Gravity is pulling the ball downward and so the force of gravity must balance the vertical component of the tension: $mg = T\cos\theta$. Moreover, since the ball is traveling in a circle, there must also be a centripetal force directed inward. The only force directed inward is the horizontal component of the tension; hence by Eq. (5.9) we require $T\sin\theta = mv^2/r$. Dividing the second equation by the first yields $\tan\theta = v^2/gr$. Taking the inverse tangent gives the required answer. If, for instance, $v = 1$ m/s and $r = 1$ m, then $\theta \approx 0.1$ rad, or roughly 6°. The answer accords with our sense of reality; we know from experience that as v increases, so does θ.

Example 5.4: Normal forces as a centripetal force. You are on a "loop-the-loop" roller coaster. What is your apparent weight at the bottom of the loop and the top?

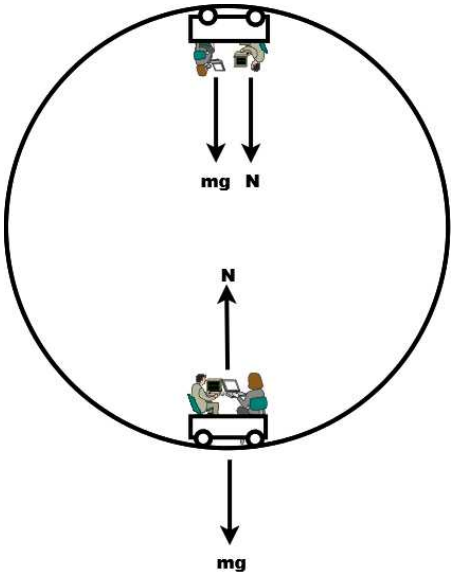


Figure 5.4

The weight you normally feel is the normal force exerted by a floor or bathroom scale on you. At the bottom of the loop-the-loop the normal force N that the car seat exerts on you is directed inward, but gravity is directed outward. The resultant of the two forces must be the centripetal force. If r is the radius of the loop and we call inward positive, then

$$\frac{mv^2}{r} = N - mg$$

or

$$N = \frac{mv^2}{r} + mg$$

Thus, at the bottom of the loop, the normal force is greater than your weight mg, and you feel heavier than usual. You know this to be true. On the other hand, at the top of the loop we have

$$\frac{mv^2}{r} = N + mg$$

or

$$N = \frac{mv^2}{r} - mg$$

and you generally feel lighter than usual. Indeed, if $v^2/r = g$, the normal force the seat exerts on you becomes zero and you fall off the ride. (**Exercise:** For what realistic values of r and v might you fall off? Why don't you? How fast would you need to be going on your favorite roller coaster to feel heavier than usual at the top of the loop?)

5. A Word About Centrifugal Force

In many daily situations, such as when a car in which we're sitting turns a corner or in a rotating amusement park ride, we feel pressed outward rather than inward. In such situations we speak colloquially of "centrifugal," or "outwardly" directed forces. Indeed, from daily experience you would conclude that it is more natural to speak of centrifugal than centripetal forces. The word "centrifuge" underscores such sentiments. Yet, the word "centrifugal" rarely appears in contemporary physics texts, and when it does the author claims that such forces do not exist; rather they are a figment of your imagination. This is complete nonsense. *Centripetal* forces appear when we describe circular motion from, say, the ground, as we have tacitly done in all examples so far in this book. *Centrifugal* forces appear when we describe the situation from inside the rotating **reference frame** itself. You cannot speak of centrifugal and centripetal forces from the same reference frame. Reference frames are discussed in more detail in Chapter 6 and a few problems in that chapter illustrate how forces appear in different reference frames.

6. Simple Harmonic Motion

Simple harmonic motion (SHM) is physicist lingo for the oscillatory, periodic motion executed by springs, pendulums and waves.[1] The first two are illustrated in Figure 5.4.

[1] Apparently the term arose because wave motion in musical instruments produce many harmonic frequencies simultaneously. "Simple" refers to the motion of a single wave.

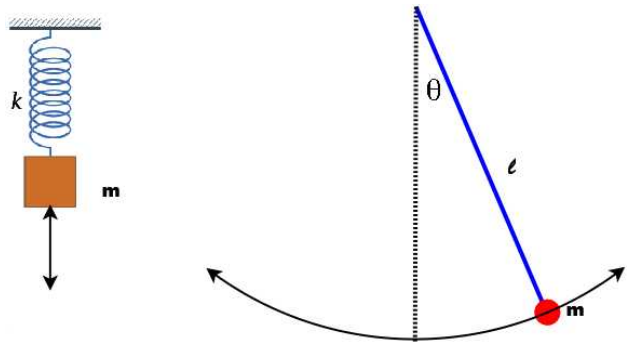


Figure 5.5

Because SHM is periodic, most of the quantities and concepts discussed for circular motion apply:

The time necessary for one roundtrip of an oscillating spring or of a swinging pendulum is termed its **period** T. The number of roundtrips per second is its **frequency** f. As stated in Eq. (5.5), $f = 1/T$. More often, one expresses frequency in terms of the **angular frequency** ω in radians per second, and the relationship between ω and f is as given by Eq. (5.7): $\omega = 2\pi f$.

A basic spring, known to cognoscenti as a **simple harmonic oscillator**, is assumed to be *massless* but have a spring constant k. The angular frequency of a mass m oscillating on such a spring is given by

$$\omega = \sqrt{\frac{k}{m}} \quad (5.11)$$

This expression is referred to as the **natural angular frequency** of the spring. To formally derive it requires calculus; however, the answer can be guessed by dimensional analysis. The single thing we know about a spring is its spring constant k, which is measured in N/m and thus has dimensions $\mathrm{M/T^2}$ (the dimensions of force/length $= ma/\ell$). Given a mass m, the *only* way we can form a frequency, which has dimensions 1/T, is by dividing k by m and taking the square root, as in Eq. (5.11). If you forget the formula for spring frequency, this is how to retrieve it. Even if the spring is oscillating in the Earth's gravitational field, Eq. (5.11) remains the only way to make a frequency from m, k and g. (**Exercise:** Try it.)

Note: The natural frequency of a spring is independent of g. It oscillates with the same frequency whether it is oscillating horizontally or vertically. As long as the spring obeys Hooke's law, the frequency is also independent of the amplitude of the oscillation.

As illustrated in Figure 5.5, a **simple pendulum** is nothing more than a mass m swinging back and forth at the end of a string of length ℓ. If the size, or **amplitude**, of the swing is $\theta \ll 1$ rad, a simple pendulum also executes SHM. The natural angular frequency of a simple pendulum swinging in the Earth's gravitational field is

$$\omega = \sqrt{\frac{g}{\ell}} \quad (5.12)$$

Once again, this result can be guessed by dimensional analysis. Because a simple pendulum is merely a mass m swinging on a string of length ℓ in Earth's gravitational field g, the frequency can depend on nothing else but these quantities (and possibly the amplitude of the swing, θ). The *only* combination of m, ℓ and g that results in a frequency is that given in Eq. (5.12). (**Exercise:** Do it.) You may also have seen long pendulums in science museums that swing very slowly compared to a ball on a string that you can hold in your hands. Your experience tells you that the longer the pendulum, the lower the frequency, as in Eq. (5.12). Note that **the frequency and hence period of a simple pendulum do *not* depend on the amplitude of the swing.** This fact has had extremely important consequences for civilization: pendulum clocks.

Example 5.5: Pendulums on the Moon. Suppose the radius of the Moon is one-quarter that of Earth. If a simple pendulum has a period T on Earth, what will be its period on the Moon, assuming that the density of the Moon and Earth are the same?

We know that on any planetary body the gravitational acceleration $g \propto M/R^2$, where M is the mass of the body and R is its radius. But $M = \rho V$, for density ρ and volume V. Furthermore $V \propto R^3$. Thus, $g \propto \rho R$; consequently, the gravitational acceleration on the Moon should be about one-quarter of Earth's. By Eqs. (5.6) and (5.12), the period is inversely proportional to $\sqrt{g}$, so we would expect the period of the pendulum on the Moon to be about $2T$. In reality, the gravitational acceleration on the Moon is about $1/6$ that of Earth, and so the pendulum's period on the Moon would be about $\sqrt{6}\,T \approx 2.5\,T$. This is typical of problems on the AP exams.

7. The Equations for Simple Harmonic Motion

SHM differs from circular motion in several obvious respects. First, a mass oscillating on a spring or a swinging pendulum is *not* moving in a circle. Also, unlike uniform circular motion, in which an object is traveling at constant speed, **the speed of an object engaged in SHM is never constant. Neither is its displacement or acceleration.**

The most important thing to remember about SHM is: **The displacement of the mass from the equilibrium position of the spring or the displacement of a pendulum trace out a sine wave as a function of time. The velocity and acceleration also trace out sine waves.** ("Sine waves" include "cosine waves.") For this reason, SHM is often referred to as **sinusoidal motion**. One can see this in Figure 5.6, where the oscillating mass produces a sine curve as a rolling piece of paper moves underneath it.

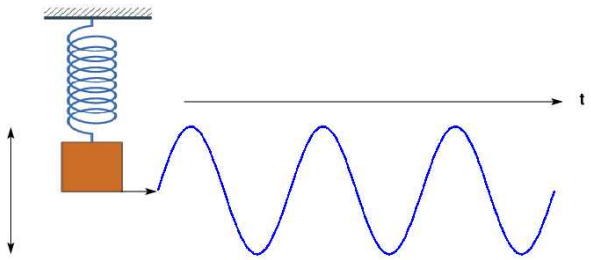


Figure 5.6

For reference, the cosine and sine as functions of an angle are plotted respectively in Figure 5.7.

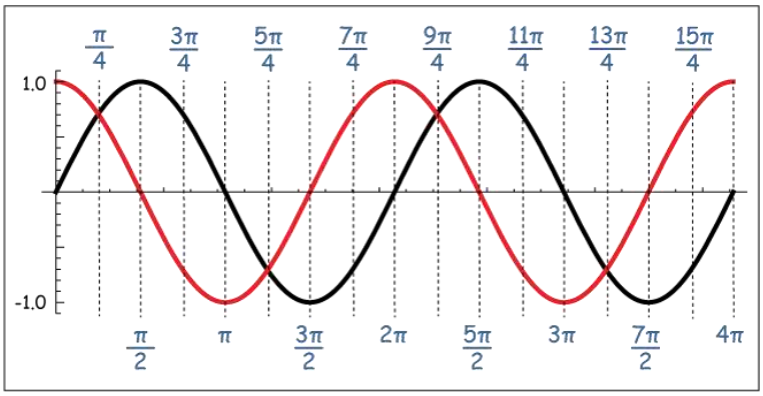


Figure 5.7

Sine and cosine require dimensionless angles (radians) for their arguments: $\sin\theta$ and $\cos\theta$. You cannot insert meters or times or anything else. You also know that these function go through one complete cycle when θ goes through an angle of 2π radians. Because a spring oscillates in time, we want to convert radians to t. From Section 4.2 we know that an object moving at ω radians per second goes through an angle θ in t seconds, as shown in Figure 5.8. Therefore, to find the object's angular position as a function of time, we write the trig functions as $\sin\omega t$ and $\cos\omega t$. Because $\theta = \omega t$, once $\theta = 2\pi$, then by definition, t equals the period T: $2\pi = \omega T$, as in Eq. (5.6). For a spring engaged in SHM, ω is the **natural frequency** given by Eq, (5.11).

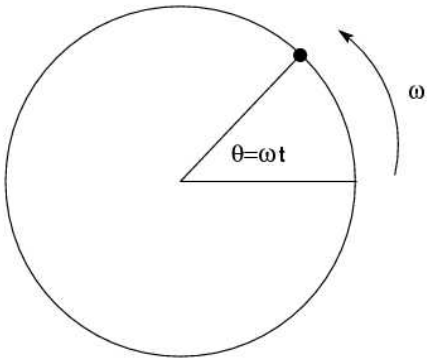


Figure 5.8

Suppose a mass on a spring is pulled a distance x_o beyond the spring's unstretched length, as shown in Figure 5.9, then released.

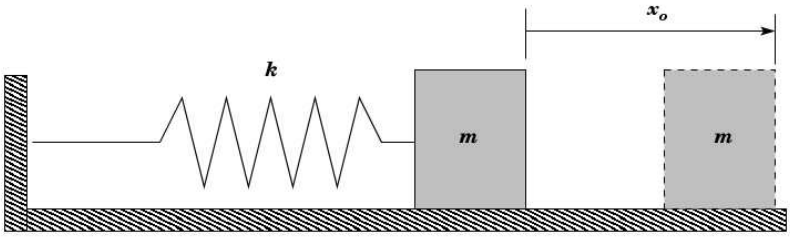


Figure 5.9

The position, velocity and acceleration of m will be described by the following equations:[2]

[2] Hooke's law requires $F = -kx$. Newton II requires $F = ma$. By definition, $a = d^2x/dt^2$. Thus, $\frac{d^2x}{dt^2} = -\frac{k}{m}x$. Students who know calculus can convince themselves by direct substitution that x given by Eq. (5.13a) or (5.14a) is a solution to this differential equation.

$$x = x_o \cos(\omega t) \quad (5.13a)$$

$$v = -\omega x_o \sin(\omega t) \quad (5.13b)$$

$$a = -\omega^2 x_o \cos(\omega t) \quad (5.13c)$$

If instead the spring is initially unstretched and the mass is whacked with a hammer, use the following equations:

$$x = x_o \sin(\omega t) \quad (5.14a)$$

$$v = \omega x_o \cos(\omega t) \quad (5.14b)$$

$$a = -\omega^2 x_o \sin(\omega t) \quad (5.14c)$$

Commentary on the SHM equations:

1) These equations apply to springs obeying Hooke's law: $F = -kx$.

2) Once again, the ω in these equations is the natural frequency given by Eq, (5.11).

3) Because $\cos(0) = 1$ at $t = 0$, Eq. (5.13a) gives $x = x_o$ at that time. The position of the mass at $t = 0$ is called the **initial position** or **initial condition**. Because $\sin(0) = 0$, Eq. (5.14a) shows that at $t = 0$ the spring is unstretched, i.e. is at its equilibrium position $x = 0$.

The only difference in behavior between $x = x_o \cos(\omega t)$ *and* $x = x_o \sin(\omega t)$ *is in the initial conditions.* The case $x = x_o \cos(\omega t)$ is similar to that in a harpsichord or guitar, where a string is first stretched, then released, whereas the case $x = x_o \sin(\omega t)$ is closer to "piano initial conditions," where a hammer strikes a string at rest and sets it into motion.

Because $|\cos\theta| \leq 1$ and $|\sin\theta| \leq 1$, $\pm x_o$ also represents the maximum displacement of the block from $x = 0$. As above, the maximum value of x in SHM is referred to as the amplitude of oscillation.

4) Eqs. (5.13) and (5.14) show that the speed of the mass reaches a maximum when the stretch is a minimum and vice versa. The maximum speed is $v_m = \omega x_o$ and the maximum magnitude of acceleration is $a_m = \omega^2 x_o$.

5) Eqs. (5.13) and (5.14) also apply to pendulums swinging through angles much less than 1 radian. (The smaller the angle, the better the equations work.) To convert the equations to those for a pendulum, merely change x to θ and a to α. One must be careful with v. The ω in these equations refers to the natural frequency of the pendulum given by Eq. (5.12), which is a *constant*. However, the linear velocity v in Eqs. (5.13b) and (5.14b) is *not* constant and neither will be the pendulum's angular velocity, which is given by v/ℓ. Usually v/ℓ is termed $\dot{\theta}$ ("theta dot") to distinguish it from ω. Thus, change v to $\dot{\theta}$ and you will be set.

Comments (2)-(4) also apply to pendulums with the above changes in notation.

Example 5.6: Energy conservation in springs. The energy of a mass oscillating on a spring must be the sum of its kinetic and potential energy: $E = \frac{1}{2}mv^2 + \frac{1}{2}kx^2$. Substituting the expressions for displacement and velocity from Eqs. (5.13), this becomes

$$E = \frac{1}{2}m\omega^2 x_o^2 \sin^2 \omega t + \frac{1}{2}kx_o^2 \cos^2 \omega t$$

But from Eq. (5.11) we know that $\omega^2 = k/m$, implying that $E = 1/2\, kx_o^2(\sin^2 \omega t + \cos^2 \omega t)$. However, by the famous Pythagorean trig identity, the expression in the parentheses equals 1, and so $E = 1/2\, kx_o^2$, always. This is just the initial potential energy of the stretched spring. Energy is conserved.

The energy diagram in Figure 5.10 clearly illustrates the situation. The total energy available to the system is the constant E_{Tot}, indicated by the horizontal line. The potential energy curve (Section 3.7 is the parabola $U = 1/2\, kx^2$ and the difference between the E_{Tot} and U at any point is the kinetic energy KE. The mass m on the spring rolls down the **potential well**. As it rolls down, KE increases while U decreases. The maximum KE, and therefore maximum velocity, occurs at the bottom of the well, where U is zero. The mass rolls up the opposite side moving ever more slowly until the **turning point**, where the horizontal line intersects the parabola. At that instant all the energy is potential and therefore the mass has zero velocity. After reaching the turning point, the mass rolls down the hill again, oscillating back and forth forever. **This is the way you should visualize springs.**

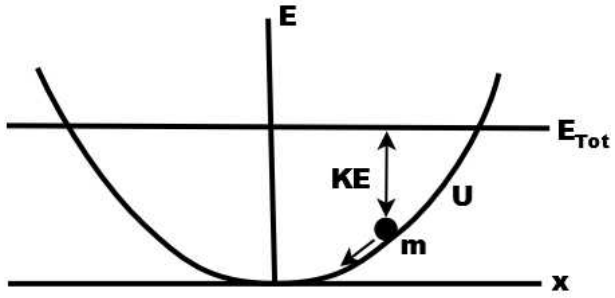


Figure 5.10

8. Wave Basics and Traveling Waves

Like Gaul, all physics is divided into three parts: billiard balls, springs and waves. To a physicist, anything that is not a billiard ball is a spring, and if it is not a spring it is a wave. Like springs, waves execute periodic motion. The chief distinction in their behavior is that whereas according to the previous equations **springs oscillate only in time, waves oscillate in both space and time.** Furthermore, a spring is a material object, a wave is not. A water wave exists in water, but the moving disturbance is not the water itself, any more than the traveling lights in Times Square proclaiming ELVIS SIGHTED IN GREAT PYRAMID are the message. The message is moving, the lights are blinking on and off in place. With waves, the medium is not the message.

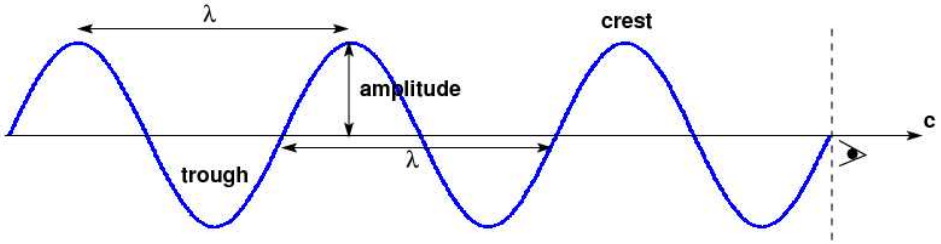


Figure 5.11

Figure 5.11 illustrates the basic properties of all waves. The distance between any two crests (or troughs) on a wave is termed the **wavelength** and is virtually always indicated by the Greek letter λ ("lambda"). The wave **amplitude** is the height of a trough or crest. (Sometimes engineers speak of "peak-to-peak" amplitude, which is the full distance between a crest and a trough, or twice the illustrated amplitude.)

Unlike springs, waves generally do not oscillate in a fixed location but propagate through space. Not surprisingly, such waves are referred to as **traveling waves**. If the wave oscillates perpendicularly to the direction of propagation, as the one drawn, it is referred to as a **transverse wave**.

The illustrated wave is to be visualized as propagating to the right with velocity c, meaning that the entire pattern shifts rigidly rightward with time. If the wave represents a water wave, a cork at the position of the dashed line will bob up and down in SHM as the wave passes that position, but will not move to the left or right. Thus, it is not the water that is propagating to the right; rather the wave is propagating through the water, just as the ELVIS message propagates through the lights on Times Square.

Alternatively, we can imagine taking a snapshot of a wave, freezing it at a given instant in time. Then as one moves along the x-axis, the frozen wave is seen to oscillate as a sine wave as x changes. This is what is meant when we say that waves oscillate in both space and time.

The time it takes for a wave to traverse one wavelength λ as observed by an observer stationed at a fixed position is by definition the wave's **period** T. Given that the wave's speed is velocity c, we must have $\lambda = cT$. The **frequency** of the wave is the number of crests passing the observer each second. As in circular motion and SHM, the frequency is $f = 1/T$. (If it takes $1/8$ second to cross one wavelength, then 8 crests pass per second.) Thus, we arrive at the fundamental relationship between velocity, frequency and wavelength that applies to all waves:

$$c = \lambda f \quad (5.15)$$

Tip: If you remember nothing else about waves, remember this.

Example 5.7: Common radio wavelengths and frequencies. Radio waves travel at the speed of light, $c = 3 \times 10^8$ m/s. A typical FM radio frequency is $f = 100$ Mhz, meaning 100 Megaherz, or equivalently, $100 \times 10^6 = 10^8$ Hz. The wavelength is $\lambda = 3 \times 10^8/10^8 = 3$ m. It is a useful entry in your number bank to remember that common FM radio waves are in the meter range. On the other hand, an AM radio station might broadcast at 530 kHz, meaning 530 kilohertz, or equivalently, 530×10^3 Hz. The wavelength is $\lambda = 3 \times 10^8/5.3 \times 10^5 \approx 6 \times 10^2$ m. This is about half a kilometer. Old "short wave" signals are typically in the range of $1 - 10$ MHz, which makes them shorter than AM radio signals but longer than FM.

Green light, on the other hand, has a wavelength of about 500 nm, meaning 500 nanometers or equivalently, $500 \times 10^{-9} = 5 \times 10^{-7}$ m. Thus, **electromagnetic waves**, which include light and radio, span an enormous range of wavelengths. **Exercise:** What is the frequency of green light? What would you say is a typical wavelength and frequency for a water wave in a swimming pool?

Because there are 2π radians between any two crests in a sine wave (Figure 5.7), the relationships $\omega = 2\pi f$ and $T = 2\pi/\omega$ still apply. Unlike a spring or pendulum, however, there is no spring constant or length of string involved. Thus, a wave has no natural frequency of oscillation, unlike those in Eqs. (5.11) and (5.12). For waves, ω is merely the angular frequency corresponding to f, whatever that happens to be.

For reference, the mathematical form of a traveling wave is generally one of the following:

$$y = A\sin(kx \pm \omega t) \;\; ; \;\; y = A\cos(kx \pm \omega t) \quad (5.16)$$

In these expressions, y is the height of the wave for a given position x and time t. A is the amplitude, or maximum height of the wave. The quantity k is called the **wave number**;

$$k \equiv 2\pi/\lambda \quad (5.17)$$

The wave number turns out to be extremely useful and you should regard it as the position analog of ω. Remember that the arguments of the trig functions must be angles. To convert angles to time we used $\omega t = \theta$; when $\theta = 2\pi$ the time t has gone through one period, T. Similarly, k converts position to angle: $kx = \theta$; when $\theta = 2\pi$, $x = \lambda$, and so the wave has gone through one wavelength.

Although we have written down Eqs. (5.16) without proof, you can see that they behave as a wave should. Pick a fixed position, $x = 0$. Then the height of the wave obeys SHM as a function of time, $y = A\sin\omega t$, as stated above. Take a snapshot of the wave at $t = 0$; then it obeys SHM as a function of x, also as stated above. Furthermore, given a function $y = f(x)$, you know from your algebra course that $y = f(x - a)$, where a is a constant, shifts the function to the right by an amount a. Thus, $y = A\sin(kx - \omega t)$ is the function $y = A\sin kx$ shifted to the right by an amount ωt. Because t is always increasing, the shift is increasing as well, and the wave travels to the right. If instead the $(+)$ appears in Eqs. (5.16), the wave travels to the left.

The wave velocity c that appears in Eq. (5.15) depends on the type of wave and the medium through which the wave is traveling. An **electromagnetic wave** (radio or light wave) in vacuum *always* travels at $c = 3 \times 10^8$ m/s. **If you haven't memorized this number, do it now.** The speed of a sound wave depends strongly on the medium and temperature. In air at standard temperature and pressure the speed of sound is about 330 m/s. **Put this number in your number bank too.** The speed of a wave on a stretched string depends on the tension in the string (how tightly it is stretched), on the string's mass M and on its length L. It is given by

$$c = \sqrt{\frac{TL}{M}} = \sqrt{\frac{T}{\mu}} \qquad (5.18)$$

where μ is the mass per unit length of the string. We do not derive this result, but dimensionally the expression in Eq. (5.18) is the only velocity that can be formed from a tension, a mass and a length. (**Exercise:** Convince yourself of this.)

9. Standing Waves and Harmonics

Stringed instruments, such as a violin, produce waves that do not travel in the same way as those described by Eqs. (5.16). *The pegs on a violin provide a physical constraint that force the ends of the string to remain stationary*. For this reason, the waves on a violin do not travel; waves that don't propagate are referred to as **standing waves**, or sometimes as **stationary waves**.

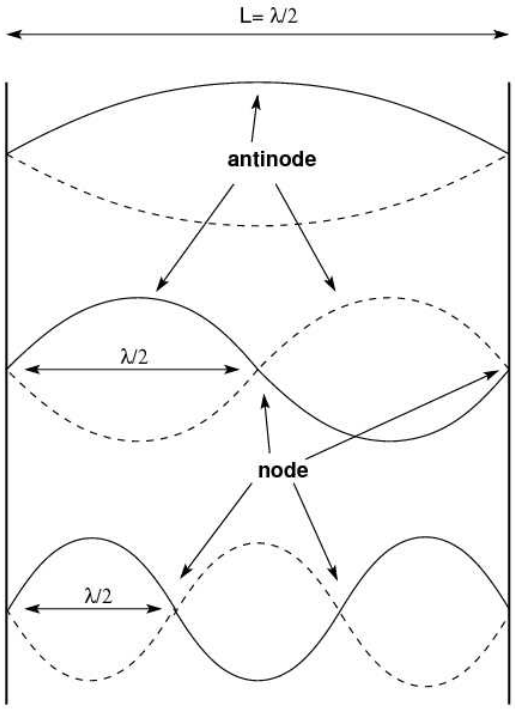


Figure 5.12

Figure 5.12 shows several possible standing waves between pegs of a violin, piano or guitar. As a function of time, any point on the string will oscillate up and down, between the solid and dashed lines. Points on the wave that remain stationary are referred to as **nodes**. Points of maximum displacement are referred to as **antinodes**. Notice that the nodes in these figures are always spaced by $\lambda/2$, half a wavelength. If L is the distance between the pegs, *the condition that the endpoints remain fixed implies that the longest possible wave that can fit between the pegs has a wavelength such that* $L = \lambda/2$. This is illustrated in the topmost drawing. The longest wave, and hence lowest frequency, is termed the **fundamental**, or **first harmonic**. The next longest wave that can be fit in (middle drawing) is termed the **second harmonic** (or first overtone) and so forth.

The allowed wavelengths that can be fit between the pegs can be found as follows: Since these are standing waves, the nodes do not change position with time. Judging from Figure 5.12, the height of the wave at any point is given by $y = A \sin kx$, where x is the distance from the left end. (This is just Eq. (5.16), where we have set $t = 0$ because the equation must true at any time t.) By definition, a node occurs when $y = 0$, and so the condition for any node is

$$A \sin kx = 0 \qquad (5.19)$$

The sine function is zero when kx is an integral multiple of π, and so the requirement for a node becomes

$$kx_n = n\pi$$

Here, n is a nonnegative integer and x_n refers to the position of the *nth* node. With the definition of k from Eq. (5.17), this becomes

$$x_n = \frac{n\lambda}{2} \qquad n = 0, 1, 2, 3 \ldots$$

In every case, the boundary condition that the string remains stationary at L requires there to be a node at $x = L$. Thus, we can also write $L = n\lambda/2$, or

$$\lambda_n = \frac{2L}{n} \qquad (5.20)$$

where now $n \geq 1$. When $n = 1$, we have $\lambda = 2L$, the fundamental, as in the top drawing of Figure 5.12. When $n = 2$, $\lambda = L$, in accord with the middle drawing, and so on. Eq. (5.20) gives the allowed wavelengths that can fit in the "box" formed by two pegs. The fact that not every wavelength is allowed means that we have **quantized** the system (see Section 12.6).

Example 5.8: Violin strings. What tension should a violin string be put under to produce A 440? This is a good problem to exercise one's physical intuition; otherwise we are not given enough information to solve it. Let's assume that the violin string is producing the fundamental. Then in Eq. (5.20), $n = 1$ and the wavelength $\lambda = 2L$, which with the basic relationship between frequency and wavelength, Eq. (5.15), can be rewritten as $c = 2fL$. From Eq. (5.18) this becomes $\sqrt{\frac{TL}{M}} = 2fL$, or $T = 4f^2ML$. The only thing we know, however, is that $f = 440$ Hz, so we must make some guesses. If you have played with your high school orchestra, or at least seen a violin, you might guess that $L \approx 30$ cm. The mass of the string is harder to guess but a good number for your number bank is that the mass of a penny is 2.5 g. Presumably, the mass of a violin string must be smaller than that. Let's guess 1 gm. Plugging in gives $T \approx 232$ N. This is a factor of $2 - 3$ higher than the tensions in real violin strings. The obvious thing to do is weigh a real violin string.

10. Sound Waves

Sound waves have a wavelength and a frequency and can be treated by the mathematics of the previous sections. Physically, however, sound waves differ from electromagnetic waves or waves on a string. A sound wave traveling through air alternately compresses and rarefies the gas through which it moves, creating alternating regions of high and low pressure that vary sinusoidally along the direction of wave propagation (Figure 5.13). Because the air molecules are being displaced from their undisturbed positions to form the regions of high and low pressure, their displacement also varies sinusoidally. One can therefore view sound waves as either pressure waves or displacement waves. Waves whose disturbance oscillates in the direction of propagation are referred to as **longitudinal waves**. If you have ever sent a wave down a slinky, you have seen a longitudinal wave in action. (In Figure 5.13, the drawn sine curve illustrates the amplitude of the compression, or change in pressure, although in air the actual molecules are not oscillating vertically; rather they are bunching and rarefying horizontally.)

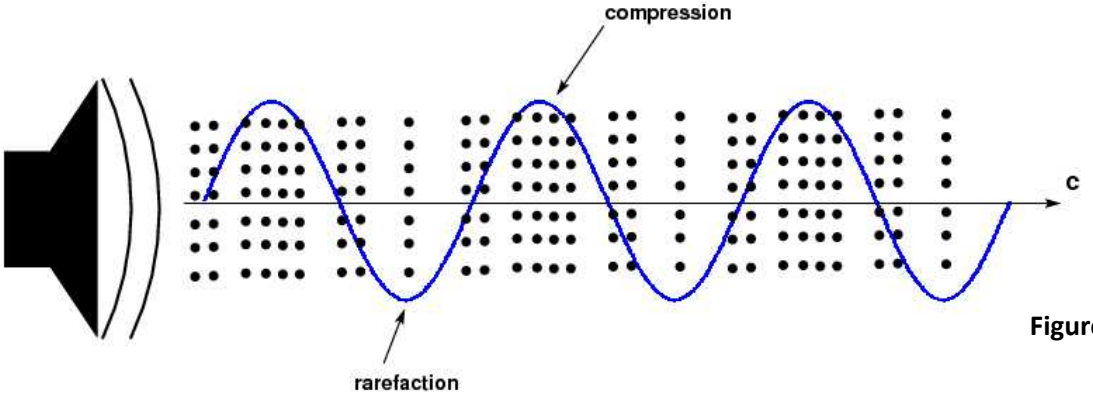


Figure 5.13

Just as standing transverse waves can be set up on pegged strings, standing longitudinal waves can be set up in tubes or musical instruments. As in the case of strings, the nodes are spaced $\lambda/2$ apart. However, in the case of sound waves in a tube, one must distinguish two sorts of nodes, **pressure nodes** and **displacement nodes**. Consider a standing sound wave oscillating in a tube closed at both ends. In analogy to the violin string pegged at both ends, the walls provide boundary conditions. More precisely, the gas molecules cannot move at the tube walls. **A stopped end of a tube represents a displacement node**. But that is also where the molecules have bunched up and is where the pressure is highest. Thus, **a stopped end of a tube represents a pressure antinode.** In contrast, an open end of a tube must be at normal atmospheric pressure. Therefore, **an open end of a tube is a pressure node.** But the air at an open end is also free to move. Therefore, **an open end of a tube is a displacement antinode.**

In tubes and wind instruments, displacement nodes are pressure antinodes and vice versa.

Certain wind instruments, such as the oboe, behave like a tube open at both ends. The clarinet, however, behaves like a tube closed at one end and open at the other. The first three allowed harmonics for the three possibilities for tubes are shown schematically in Figure 5.14. In each case, there must be a pressure node at an open end (blue/bold) and a displacement antinode (red) at the same place. There must also be a pressure antinode and a displacement node at a closed end. (In order not to be confused when looking at these diagrams, it is best to concentrate on either the displacement oscillations or the pressure oscillations separately.)

For the tube closed at both ends (closed-closed) and the tube open at both ends (open-open) one can find the allowed harmonics in the same way as we found those for the standing waves on a violin. Focusing attention on the displacement curve, we see that it is a sine function with the same boundary conditions as Eq. (5.19). Thus, the allowed **modes** are the same:

$$\lambda_n = \frac{2L}{n} \quad n = 1,2,3 \ldots \quad (5.21)$$

$$closed - closed \; or \; open - open \; tubes$$

(**Mode** is a term encountered frequently in physics to describe a given pattern of vibration, including harmonics.)

For the closed-open tube, one sees from the bottom diagram in Figure 5.14 that L can be only $\frac{\lambda}{4}, \frac{3\lambda}{4}, \frac{5\lambda}{4}, \ldots$ Thus, the condition on allowed wavelengths is

$$\lambda_n = \frac{4L}{n} \quad n = 1,3,5 \ldots \quad (5.22)$$

$$closed - open \; tube$$

(**Exercise:** Convince yourself that Eq. (5.22) is correct.)

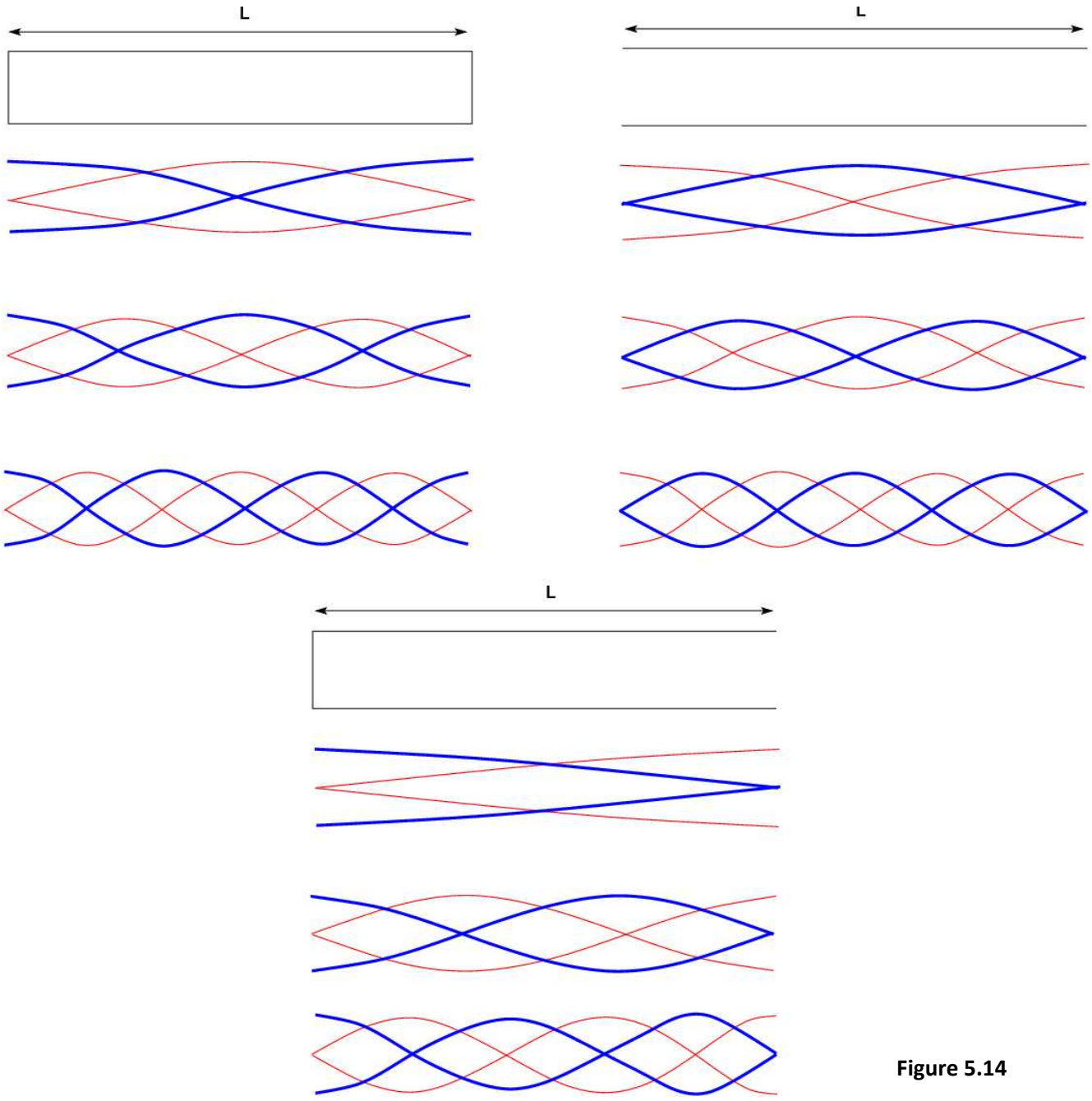


Figure 5.14

Example 5.9: Musical wavelengths. What would be the length of an open-open tube needed to produce A 440?

The speed of sound varies with temperature. From tables, at room temperature $c_s \approx 340$ m/s. The basic relationship between frequency and wavelength, Eq. (5.15), then gives $\lambda = 340/440 = 0.77$ m. If this is the fundamental, then according to Eq. (5.21) the length of the tube should be $L = 0.77/2 \approx 38$ cm. Reality check: Is this a reasonable length for a musical wind instrument?

Example 5.10: Steam whistles. Old steam whistles on trains acted like closed-open tubes. Assume that the whistle was $0.5\ m$ long and that when it was sounded, the first overtone above the fundamental was heard. What is the frequency and wavelength? If the tube were open-open, would the first overtone be higher or lower? By what factor?

According to Eq. (5.22), the fundamental is at $\lambda = 4L = 2$ m, and the first overtone is $\lambda = 4/3\ L = 2/3$ m. By the relationship $f = c/\lambda$, we have $f = 510$ Hz, slightly higher than concert C. If the tube were open-open, Eq. (5.21) says that the fundamental is $\lambda = 2L = 1$ m, and the first overtone is $\lambda = L = 0.5$ m. This wavelength is $3/4$ the previous, and so the frequency is a factor of $4/3$ higher, or 680 Hz.

11. Resonance

The above examples of frequencies and wavelengths supported by musical pipes provide a good illustration of **resonance**. If you imagine putting a small loudspeaker at the mouth of the open-open pipe in Example 5.9 and dialing in different frequencies with a tone generator, the sound will get louder at the fundamental frequency 440 Hz, then become soft again until you dial in the next allowed frequency given by Eq. (5.21), $f = 880$ Hz, the second harmonic. The harmonics supported by the pipe are also termed **natural frequencies of oscillation** or **resonant frequencies.** The system easily absorbs energy at these frequencies but not at others; scientists say that the resonant frequencies are those *excited* by the tone generator. **Resonance is the amplification of natural frequencies by excitation.**

Many physical systems exhibit resonance. We have already written down the natural frequencies for the spring and the pendulum in Eqs. (5.11) and (5.12). If one imagines tapping a mass on a spring or pushing a pendulum at its natural frequency, it is easy to believe that the amplitude of oscillations will grow steadily, but that if you push on a pendulum with some other frequency, you may even stop it. You witness this when you push a child on a swing. If you push at the swing's natural oscillation frequency, the system absorbs energy and the swing goes higher and higher; if you don't, you may end up flat on your back.

Historically, the most famous example of apparent resonance is the collapse of the Tacoma Narrows Bridge in 1940, which has become the favorite video in many physics classes. However, recent analyses indicate that the collapse was almost certainly not in fact due to resonance. Nevertheless, examples of true resonance abound. A laser relies on resonance in its optical cavity and a radio tuner provides an example of electromagnetic resonance; your selected station comes in loud and clear only when the tuner is adjusted to the same frequency and can absorb electromagnetic energy. Although the mathematics describing resonance is above the level of this text, the subject is an important one and the SAT subject test and AP exams do expect you to have a basic understanding of what it is.

12. Doppler Shift

You have undoubtedly noticed that when an ambulance or police car approaches you, the pitch of the siren increases and when it recedes from you in the distance, the pitch decreases. This is an example of the famous **Doppler shift**, after Christian Doppler, who proposed it in 1842 to explain the various colors of starlight.

Suppose a loudspeaker is emitting sound waves in the form of pulses with wavelength λ_o, as in Figure 5.15a. If you are at rest relative to the speaker, you will hear the sound waves with exactly the same wavelength, λ_o and same frequency f_o. On the other hand, suppose the speaker is moving toward you with a speed v, as in Figure 5.15b. In one wave period T the speaker will have moved a distance $d = vT$, when it gives out the next pulse. Therefore, as you can see from the figure, the wavelength of the sound you receive is $\lambda_r = \lambda_o - vT$. From the basic wave relationship Eq. (5.15), $\lambda_o = cT$, where c is the wave speed, and so we have $\lambda_r = (c - v)T = (c - v)/f_o$. But $\lambda_r = c/f_r$, which gives

$$f_r = \frac{cf_o}{c - v} = \frac{f_o}{1 - \frac{v}{c}}\ .$$

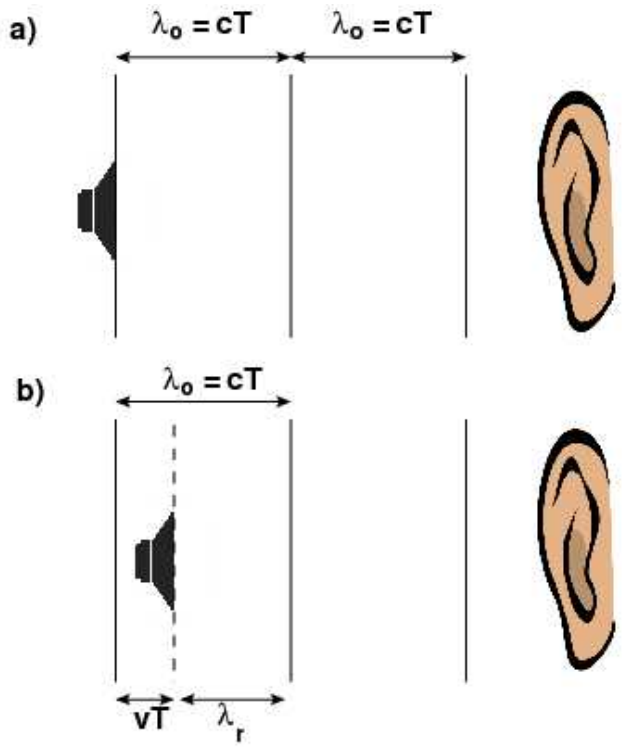


Figure 5.15

If $v \ll c$, as it is in most cases, one can use a well-known approximation[3] to write

$$f_r \approx f_o\left(1+\frac{v}{c}\right).$$

From this form you can readily see that f_r is greater than f_o, which accords with your experience that the pitch of an approaching siren is raised from its normal frequency. In the same way you can show (**Exercise**: Do it) that if the speaker is moving away from you,

$$f_r \approx f_o\left(1-\frac{v}{c}\right),$$

and thus the pitch is lowered. Both formulas can be combined into

$$f_r \approx f_o\left(1 \pm \frac{v}{c}\right) \quad (5.23)$$

Textbooks work out similar formulas if the *receiver* is moving, rather than the source, but in that case Eq. (5.23) holds *exactly*. (**Exercise:** Try it**.** Hint: When you move toward the receiver the velocity of the sound wave you hear changes to $c_r = c + v$.) Thus, when $v \ll c$ you can use it for either situation. If *both* receiver and source are moving, one can go through the same sort of argument (**Exercise**: try it) to get a more complicated formula. But again, if $v \ll c$ it is essentially the same as above:

$$f_r \approx f_o\left(1 \pm \frac{v_s}{c} \pm \frac{v_{ob}}{c}\right) \quad (5.24)$$

---

[3] The binomial expansion says that $(1+\epsilon)^n \approx 1 + n\epsilon$ when $\epsilon \ll 1$. In our case $v/c = \epsilon$.

Here v_s is the velocity of the source, v_{ob} is the velocity of the observer, $(+)$ indicates motion toward the other, $(-)$ indicates motion away from the other. Eq. (5.24) makes sense: you'd expect a higher frequency if the source and observer are moving toward each other, and a lower frequency if they are moving away from each other.

Example 5.11: Sound the trumpets. The first demonstration of the Doppler effect was carried out in 1845 by Buys Baillot when he placed a trumpeter on an open train car pulled by a locomotive. Another trumpeter stood at rest in the station, sounding the same note. Citizens from Utrecht were to judge whether the Doppler effect existed. Was it possible?

Apparently the speed of the train was $40\ \text{mph}$, or about $18\ \text{m/s}$. Taking the speed of sound to be $340\ \text{m/s}$, $v/c \approx 0.0529$. The trumpeters were sounding concert G, which has a (modern) pitch of 392 Hz. From Eq. (5.23) the spectators would have heard a frequency $f_r = 413$ Hz from the moving trumpeter as the train approached, if he was not drowned out by the locomotive. Each half-step on a musical scale is obtained by multiplying the previous note by $\sqrt[12]{2}$. Multiplying 392 Hz by $\sqrt[12]{2}$ gives 415 Hz, close to f_r. Spectators would have thus heard from the train a concert G#. Even those with a moderately good sense of pitch would have realized that the two trumpeters were playing out of tune. The demonstration therefore appears possible.

13. Superposition and Interference

In Section 9 we discussed standing waves. Standing waves, like those illustrated in Figure 5.12, can be viewed as a **superposition** of right-traveling and left-traveling waves.

Principle of superposition

When two waves pass through each other, each wave remains undisturbed but they produce a resultant wave whose amplitude is the algebraic sum of the two incident waves.

Superposition gives rise to **interference**. When two waves are superposed, they can add such that in some places the total amplitude is greater than that of each wave, and in other places "cancel each other out." The former is termed **constructive interference**; the latter is termed **destructive interference**.

Pure constructive interference takes place when the crest of the first wave falls directly atop the crest of the second wave. Pure destructive interference takes place when the crest of the first wave falls directly atop the trough of the second wave.

Figure 5.16 illustrates these situations. In the top drawing, the crests of the identical blue and green waves are shifted by a full wavelength, and so the two individual waves add to make a wave of twice the original amplitudes (dashed red wave). In the bottom drawing, the blue and green waves are shifted by one-half wavelength and so exactly cancel out to make zero disturbance (red dashed line).

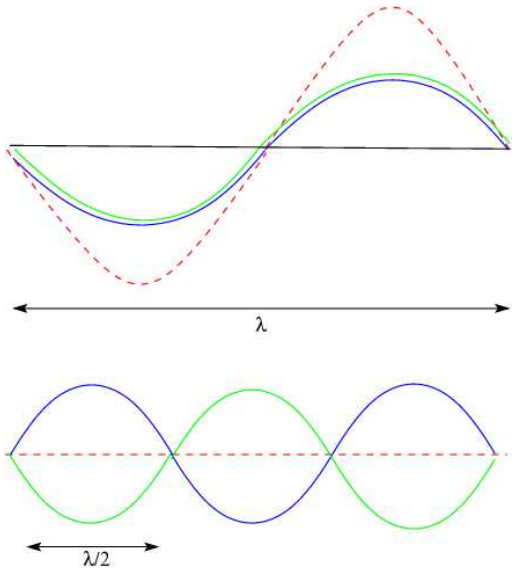


Figure 5.16

When two waves interfere constructively, we say they are **in phase**; when they interfere destructively, we say they are **out of phase**. Two waves that are perfectly in phase must have their crests an integral number of wavelengths apart, as in Figure 5.16. Two waves that are perfectly out of phase must have their crests or troughs separated by half a wavelength or an odd-integer number of half-wavelengths (Why?).

Exercise: Carefully add the cosine and sine waves in Figure 5.7 and observe where the interference is constructive (amplitudes add) and where it is destructive (amplitudes subtract).

It is supremely important to realize that billiard balls and springs do not interfere; waves do. When you cannot hear your favorite radio station because another station is interfering with it, you are experiencing exactly this property of waves: the waves from the two stations are combining in such a way that you cannot make anything out.

The ability to undergo superposition and interference is the defining property of waves.

As mentioned above, standing waves may be viewed as the result of superposition of traveling waves.

Example 5.12: Standing waves from traveling waves. Suppose we have a left-moving sine wave interfering with a right-moving sine wave of equal amplitude. According to the principle of superposition and the first of Eqs. (5.16), the amplitude of the combined wave should be

$$y = A\sin(kx + \omega t) + A\sin(kx - \omega t)$$

With the trig identity $\sin(a \pm b) = \sin a \cos b \pm \cos a \sin b$, this expression can be rewritten as

$$y = 2A \sin kx \cos \omega t$$

(**Exercise:** Do it.)

Notice that the first factor depends only on x and the second factor depends only on t. Thus, regardless of the value of t, y will equal zero when $2A \sin kx$ equals zero. This is the same condition for a node that was stated in Eq. (5.19) earlier. In this case the two traveling waves are interfering in such a way as to produce a standing wave with nodes in fixed positions.

Further aspects of superposition and interference will be explored in the problems. Interference and diffraction of light waves will be covered in Sections 11.11-11.12.

CIRCULAR MOTION, SIMPLE HARMONIC MOTION AND WAVES PROBLEMS

Problems 1-20 are at the subject test level. The remainder are at the AP level.

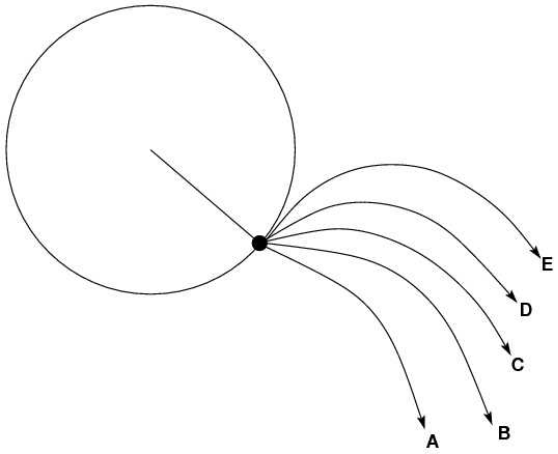
1. A satellite in orbit around the Earth at a distance $r = 6800$ km from the Earth's center, completes 3 complete orbits in 4.5 hours. Its angular speed is

   (A) 4.5×10^{-5} rad/s
   (B) 7.3×10^{-3} rad/s
   (C) 1.16×10^{-3} rad/s
   (D) 4.5×10^{-2} rad/s
   (E) 0.116 rad/s

2. A grinding wheel is sped up to 600 revolutions per minute, then the switch is turned off. If the angular acceleration is -2 rad/s^2, how long will it take the wheel to stop rotating?

   (A) 0.314 s
   (B) 3.14 s
   (C) 31.4 s
   (D) 6.28 s
   (E) 60.28 s

3. You are whirling a ball around on a string, counterclockwise in a vertical circle. At the point shown below the string is suddenly cut.



   Which of the indicated paths most closely resembles the path that the ball will follow before hitting the ground?

   (A) A
   (B) B
   (C) C
   (D) D
   (E) E

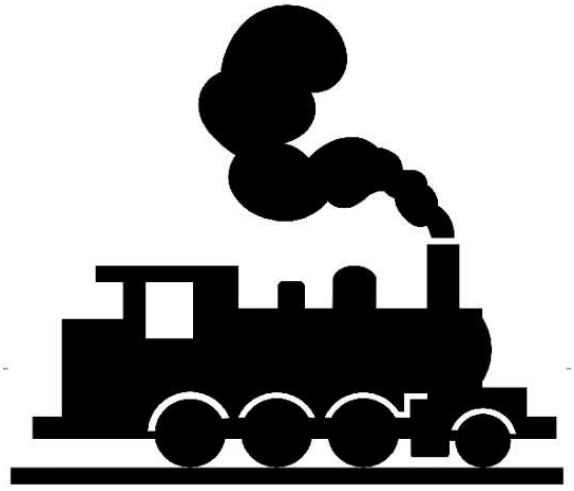
4. A helium-neon laser produces red light at a wavelength of 633 nm. The frequency f and angular frequency ω of the light are most nearly

   (A) $f = 4.7 \times 10^{14}$ Hz ; $\omega = 3.0 \times 10^{15}$ rad/s
   (B) $f = 4.7 \times 10^{14}$ Hz ; $\omega = 7.5 \times 10^{13}$ rad/s
   (C) $f = 3.0 \times 10^{15}$ Hz ; $\omega = 4.7 \times 10^{14}$ rad/s
   (D) $f = 4.7 \times 10^{-14}$ Hz ; $\omega = 3.0 \times 10^{-15}$ rad/s
   (E) $f = 4.7 \times 10^{14}$ rad/s ; $\omega = 7.5 \times 10^{13}$ Hz

5. Rank the following from longest to shortest wavelength: visible light, FM radio signals, x-rays, AM radio signals, microwaves.

   (A) AM, FM, light, microwaves, x-rays
   (B) Light, x-rays, AM, FM, microwaves
   (C) Microwaves, light, x-rays, AM, FM
   (D) FM, microwaves, light, x-rays, AM
   (E) AM, FM, microwaves, light, x-rays

Questions 6 - 8 refer to the following figure to the right.



6. The locomotive shown has small wheels with radius r and large wheels with radius R. The locomotive attempts to move to the right but the tracks are icy and the wheels spin in place. Consider a point P on the edge of each wheel. If the angular acceleration of the wheels is α, then after a time t, which of the following statements is true?

   i. All the wheels have the same angular speed about their centers.
   ii. The points P on the edge of the wheel will have traveled the same distance on all the wheels.
   iii. The points P will have traveled through the same angle.
   iv. The points P have the same linear speed.

   (A) i and ii
   (B) ii and iii
   (C) iii and iv
   (D) i and iii
   (E) ii and iv

7. The ice remains on the track. If the angular acceleration of the wheels is $0.2\ \text{rad/s}^2$ and after a time t the wheels have rotated through an angle $\phi = 360$ rad, what is the angular speed at that time?

   (A) 2 rad/s
   (B) 3 rad/s
   (C) 6 rad/s
   (D) 12 rad/s
   (E) 18 rad/s

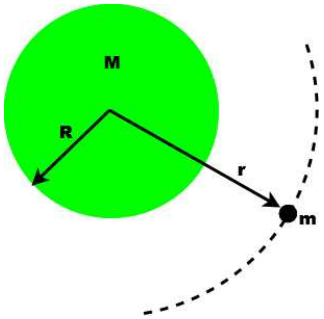
8. The ice has melted and the locomotive wheels no longer slip on the track. The locomotive begins to move with acceleration $a = 0.1\ \text{m/s}^2$. The radius of the small wheels is $r = 0.2\ \text{m}$ and the radius of the large wheels is $R = 0.4\ \text{m}$. After 100 s, which of the following statements are true? (More than one answer is permitted.)

   (A) Points P on the edges of all the wheels have the same linear speed.
   (B) All the wheels have the same angular speed around their centers.
   (C) All the wheels have performed the same number of revolutions.
   (D) The large wheels have performed half the number of revolutions as the small wheels.
   (E) The large wheels have performed twice the number of revolutions as the small wheels.

9. A car travels in a circle on a level road. It is able to do this because

   (A) the car follows the curve of the road.
   (B) friction between the tires and the road provides a centripetal force that allows the car to do this.
   (C) gravity keeps the car on the road.
   (D) the road exerts a normal force on the tires.
   (E) (B) and (D)

Questions 10 - 13 refer to the following figure.



10. A satellite of mass m is in a circular orbit around planet Alpha of mass M and radius R. The radius of the orbit is r, as shown. To be sure to decrease the speed of the satellite, Khan the Conqueror could simultaneously

    (A) increase M; decrease r.
    (B) decrease M; increase r.
    (C) increase M; leave r unchanged.
    (D) leave M unchanged; decrease R.
    (E) leave M unchanged; increase R.

11. If Khan doubles M and replaces the satellite of mass m by a satellite of mass $2m$ at the same orbital radius r as the original satellite, then

    (A) satellite $2m$ will have half the orbital speed of satellite m.
    (B) satellite $2m$ will have twice the orbital speed of satellite m.
    (C) satellite $2m$ will have the same orbital speed of satellite m.
    (D) satellite $2m$ will have $\sqrt{2}$ times the orbital speed of satellite m.
    (E) satellite $2m$ will have $1/\sqrt{2}$ times the orbital speed of satellite m.

12. If Khan were to move the satellite in Problem 10 to an orbit with orbital radius $2r$, its angular speed would

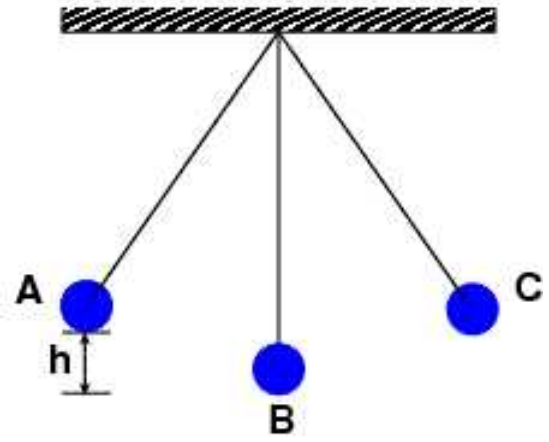
    (A) double.
    (B) halve.
    (C) stay the same.
    (D) go up by $2\sqrt{2}$.
    (E) go down by $2\sqrt{2}$.

13. If Khan moved the satellite in Problem 10 to an orbit with orbital radius $2r$, v would

    (A) go down by $2\sqrt{2}$.
    (B) go down by $\sqrt{2}$.
    (C) stay the same.
    (D) go up by $\sqrt{2}$.
    (E) go up by $2\sqrt{2}$

Questions 14 - 17 refer to the following information.

A simple pendulum consists of a mass m attached by a string of length ℓ to the ceiling. It oscillates without friction between points A and C, rising to a height h above its equilibrium position B.



14. If $m = 0.2$ kg and the speed of the bob at point B is 3 m/s, h is most nearly

    (A) 0.15 m
    (B) 0.25 m
    (C) 0.45 m
    (D) 0.65 m
    (E) 0.85 m

15. What is the total energy at point C?

    (A) 0.5 J
    (B) 0.9 J
    (C) 1.0 J
    (D) 1.8 J
    (E) 2.2 J

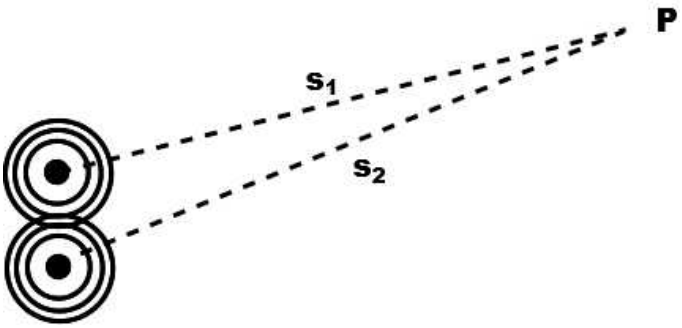
16. What is the momentum at point C?

(A) $0\ \text{kg m/s}$
(B) $0.45\ \text{kg m/s}$
(C) $0.9\ \text{kg m/s}$
(D) $1.8\ \text{kg m/s}$
(E) $2.2\ \text{kg m/s}$

17. If the length of the pendulum is increased by a factor of four,

(A) the period and frequency stay the same.
(B) the period and frequency go up by a factor of two.
(C) the period doubles and the frequency halves.
(D) the period goes up by a factor of four and the frequency goes down by a factor of four.
(E) the period and frequency go down by a factor of two.

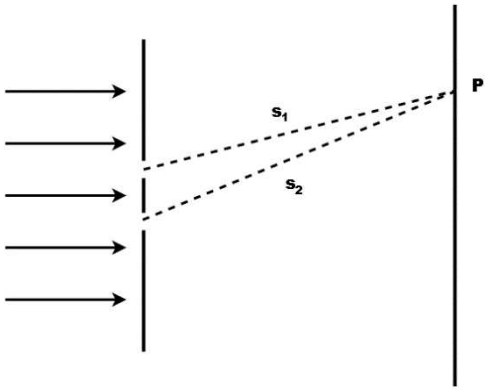
18. Two rods oscillate up and down in a tank of water, sending out ripples with wavelength λ between wave crests as shown.



If at a point P there is no disturbance in the water, one can be sure that $s_2 - s_1$, the difference in distances between P and the two sources, is

(A) $n\lambda$, where n is an integer.
(B) $\lambda/2$.
(C) zero.
(D) $(n \pm 1/2)\lambda$, where n is an integer.
(E) $n\lambda/2$, where n is an integer.

19. A uniform beam of light of a single frequency is incident perpendicularly on a screen in which there are two narrow slits, as shown. On a distant wall appear a series of light and dark stripes.



If at a point P on the wall appears as a bright stripe, one can be sure that $s_2 - s_1$, the difference in distances between P and the two slits, is

(A) $n\lambda$, where n is an integer.
(B) $\lambda/2$.
(C) zero.
(D) $(n \pm 1/2)\lambda$, where n is an integer.
(E) $n\lambda/2$, where n is an integer.

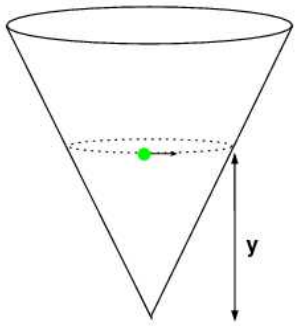
20. Two identical police sirens emit sound waves of the same frequency. Standing on a street corner, you hear siren A at a higher pitch than siren B. Which statement must be true?

(A) Siren A is moving toward you and siren B is moving away from you.
(B) Siren A is at rest and siren B is moving away from you.
(C) Siren A is moving away from you more slowly than siren B is moving away from you.
(D) Siren A is moving toward you faster than siren B is moving toward you.
(E) None of the above

21. You are on an amusement park ride in the form of a large, vertical cylinder that spins around, pinning you to the wall. After you are pinned to the wall, the floor lowers, leaving you pinned without touching the floor. The condition that you not fall is that

    (A) the centripetal force the wall exerts on you must equal or exceed the force of gravity on you.
    (B) your centripetal acceleration must equal or exceed the acceleration of gravity.
    (C) the normal force exerted by the wall on you must equal or exceed the force of gravity on you.
    (D) the frictional force exerted on you by the wall must equal or exceed the force of gravity on you.
    (E) (A) and (D)

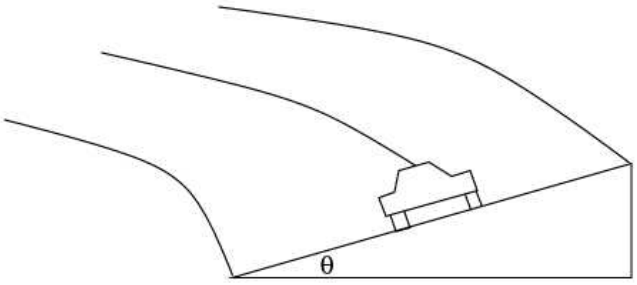
22. A marble rolls around in a circle inside a hollow cone at a constant speed and at a constant height y above the cone's apex, as shown. The radius of the marble's orbit is r.



    If y is increased, in order to continue orbiting in a horizontal circle, the marble must

    (A) roll faster.
    (B) roll slower.
    (C) roll at the same rate.
    (D) What happens depends on the coefficient of friction between the marble and the cone.
    (E) What happens depends on the width of the cone.

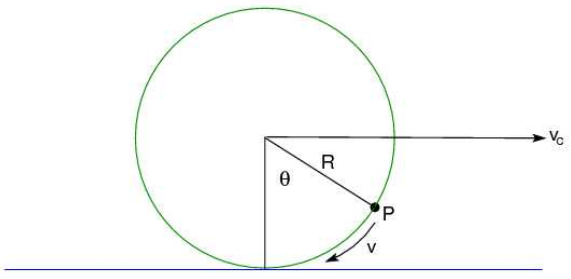
23. A car makes a perfect bank on a road tilted at an angle θ, as shown. A perfect bank is one in which the car moves neither up nor down without relying on friction.



    The condition for a perfect bank is

    (A) $mg = N$, where N is the normal force exerted by the road on the car.
    (B) $(mv^2)/r = N$, where r is the radius of the curve on which the car is traveling.
    (C) $\tan\theta = v^2/rg$.
    (D) $\cot\theta = v^2/rg$.
    (E) $rg = v^2$.

24. A wheel of radius R rolls along a horizontal surface at a constant velocity without slipping. The speed of the center of the wheel is v_c. A point P on the edge of the wheel rolls through an angle θ.



    What is the tangential speed v of the point P compared to v_c, the horizontal speed of the center of the wheel?

    (A) $v = v_c$
    (B) $v > v_c$
    (C) $v < v_c$
    (D) The answer depends on R.
    (E) The answer depends on θ.

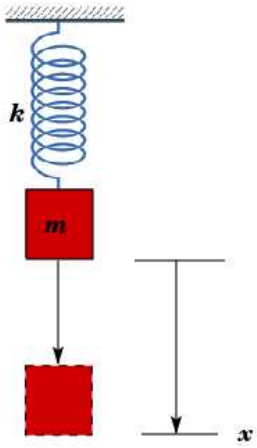
25. The total energy of a satellite in a circular orbit around any planet is

$$E = KE + U = 1/2\, mv^2 - GMm/r,$$

where the first term is the satellite's kinetic energy and the second term is its gravitational potential energy. For this system, which of the following statements are correct? (More than one answer permitted.)

(A) $KE = -U/2$
(B) $E = -KE$
(C) $E = -U$
(D) $E = +U/2$
(E) $E = -U/2$

Questions 26 - 27 refer to the following diagram:



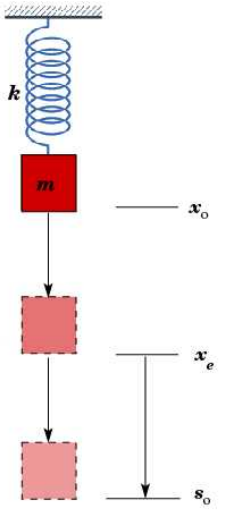
26. A mass m is attached to a spring, which is hung vertically from a ceiling. The spring stretches an amount x from its unstretched length before coming to rest. The spring constant k of the spring is equal to

(A) x/mg
(B) x^2/mg
(C) mg/x
(D) xmg
(E) None of the above

27. The period of oscillation of the spring in the previous problem will depend on

(A) m, k and g.
(B) m, k and x.
(C) m and g.
(D) g and x.
(E) m and k.

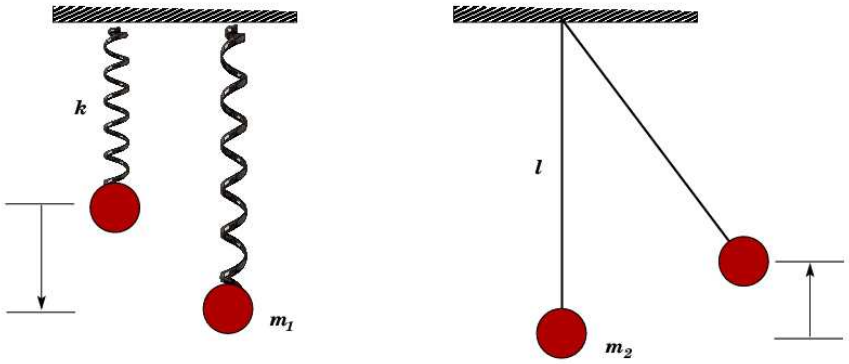
28. The mass in Problem 26 has been hanging at rest at the equilibrium position, which we now label x_e. It is then stretched a further distance s_o, as shown below, and released.



If the original, unstretched length of the spring is called x_o, which of the following is true? (More than one answer is permitted.)

(A) The spring will oscillate around x_o.
(B) The spring will oscillate around x_e.
(C) The spring will oscillate around s_o.
(D) The spring will oscillate up to its unstretched position x_o.
(E) The amplitude of the oscillation will be s_o.

Questions 29 - 30 refer to the following diagram:



29. A spherical mass m_1 is attached to a spring with a spring constant k and a spherical mass m_2 is attached to a simple pendulum of length ℓ. The spring is stretched a certain distance below its equilibrium position and released. The pendulum is raised a certain distance above its equilibrium position and released. Assuming that both spring and pendulum execute SHM, which of the following statements is true for both systems? (More than one answer permitted.)

(A) The maximum kinetic energy is attained at the equilibrium position.
(B) The maximum energy is reached only at the equilibrium position.
(C) The maximum gravitational potential energy is attained at the equilibrium position.
(D) The maximum gravitational potential energy is attained at the point of release.
(E) The minimum potential energy is reached at the equilibrium position.

30. Suppose in problem 29, the period of the spring is observed to be twice the period of the pendulum. What is the spring constant?

(A) x/mg
(B) x^2/mg
(C) $mg/4\ell$
(D) $2mg/\ell$
(E) mg/ℓ

31. A mass of 20 gm is attached to a massless spring with spring constant k. The system oscillates according to the equation $x = 4\cos 5t$, where x is measured in centimeters. If x_o is the mass' initial position, v_o is its initial velocity and f is the frequency, which of the following is true?

(A) $x_o = 4$ cm; $v_o = 0$; $k = 500$ N/m; $f = 5$ Hz.
(B) $x_o = 4$ cm; $v_o = 0$; $k = 500$ N/m; $f = 0.79$ Hz.
(C) $x_o = 0$ cm; $v_o = 4$ cm/s; $k = 500$ N/m; $f = 0.79$ Hz.
(D) $x_o = 4$ cm; $v_o = 0$; $k = 5$ N/m; $f = 0.79$ Hz.
(E) $x_o = 0$ cm; $v_o = 4$ cm/s; $k = 5$ N/m; $f = 1/5$ Hz.

32. Consider the same system as in the previous problem. If E is the total energy of the system, KE is the kinetic energy and PE is the potential energy, after π seconds, which of the following statements is true?

(A) $E = 0$ J ; $KE = 0$ J ; $PE = 0.4$ J
(B) $E = 0$ J ; $KE = 0$ J ; $PE = 0$ J
(C) $E = 0.4$ J ; $KE = 0$ J ; $PE = 0.4$ J
(D) $E = 0.4$ J ; $KE = 0.4$ J ; $PE = 0$ J
(E) $E = 0.4$ J ; $KE = 0.2$ J ; $PE = 0.2$ J

33. Consider the same system as in the previous two problems. What is the shortest time after releasing the mass when $3/4$ of the energy of the system is PE?

(A) 3.14 s
(B) $0.210\ s$
(C) 0.105 s
(D) 6.28 s
(E) 0 s

34. A wave has the form $y = 2\sin(x - t)$. If λ is the wavelength of the wave in m, f the frequency in Hz and c the wave speed in m/s, the given expression represents

(A) a wave traveling to the left with $\lambda = 1$ m, $f = 1$ Hz, and $c = 2$ m/s.
(B) a wave traveling to the right with $\lambda = 1$ m, $f = 1$ Hz, and $c = 2$ m/s.
(C) a wave traveling to the left with $\lambda = 2\pi$ m, $f = 1/2\pi$ Hz, and $c = 2$ m/s.
(D) a wave traveling to the right with $\lambda = 2\pi$ m, $f = 1/2\pi$ Hz, and $c = 2$ m/s.
(E) a wave traveling to the right with $\lambda = 2\pi$ m, $f = 1/2\pi$ Hz, and $c = 1$ m/s.

35. You are rocking a small rowboat at a rate of 12 oscillations in 36 seconds. The waves you produce travel to another boat 90 m away in 30 s and begin to rock it with an amplitude of 15 cm. Your friends aboard that boat decide to write an equation for the wave you produced. A plausible equation for the displacement y of their boat is:

(A) $y = 0.15\sin(2.1x - 0.69t)$
(B) $y = 15\sin(0.69x - 2.1t)$
(C) $y = 0.15\sin(1.4x + 2.1t)$
(D) $y = 0.15\sin(0.69x - 2.1t)$
(E) $y = 0.15\sin(1.4x - 0.48t)$

36. The amplitude of a guitar string as a function of the distance from the left end, x, and time, t, is given by $y = 0.003\sin(4x)\cos(2080t)$, where x and y are in meters. What is the length of the string, ℓ, and the pitch, f, that it produces?

(A) $\ell = 0.78$ m ; $f = 662$ Hz
(B) $\ell = 0.78$ m ; $f = 331$ Hz
(C) $\ell = 1.56$ m ; $f = 331$ Hz
(D) $\ell = 2$ m ; $f = 33$ Hz
(E) $\ell = 1.56$ m ; $f = 33$ Hz

37. Two musical pipes, one opened at both ends and the other closed at one end and open at the other, are sounding the same pitch. Let L_1 be the length of the former tube and let L_2 be the length of the latter. If we assume that they are both sounding the fundamental, then

(A) $L_1 = 2L_2$
(B) $L_2 = 2L_1$
(C) $L_1 = L_2$
(D) $L_1 = 3/4L_2$
(E) $L_2 = 3/4L_1$

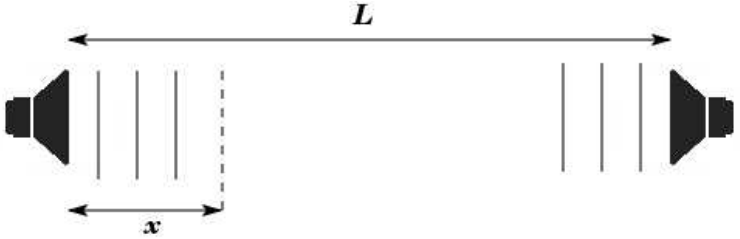
38. Two musical pipes, one open at both ends and the other closed at one end and open at the other, are of the same length. The second pipe is sounding a pitch a perfect fifth higher than the first. (A perfect fifth means that the frequencies are in the ratio 3:2.) If n_1 represents the allowed harmonics of the first pipe and n_2 the second, possible values of n_1 and n_2 are

(A) $n_1 = 1$; $n_2 = 1$
(B) $n_1 = 3$; $n_2 = 1$
(C) $n_1 = 3$; $n_2 = 15$
(D) $n_1 = 3$; $n_2 = 9$
(E) $n_1 = 2$; $n_2 = 6$

39. Two piano strings have the same mass per unit length, but string two is twice the length of string one. If string one sounds a perfect fifth above string two, then the tension in string one relative to string two is

(A) $T_1/T_2 = 9/16$
(B) $T_1/T_2 = 3/4$
(C) $T_1/T_2 = 1$
(D) $T_1/T_2 = 4/3$
(E) $T_1/T_2 = 3/2$

40. Two identical speakers facing each other across a separation L are pulsing simultaneously and emitting identical sound waves of wavelength λ. An observer moves back and forth between the speakers.



The condition that the observer hears a maximum sound volume in terms of the distance x from the left speaker and L is given by

(A) $x = n\lambda$, where n is an integer.
(B) $x = L - n\lambda$
(C) $x = (L - n\lambda)/2$
(D) $x = (n \pm 1/2)\lambda$, where n is an integer.
(E) $x = L/2 - n\lambda$

41. A police car is traveling at $34\ \mathrm{m/s}$ south with the siren on. As the police car recedes, a person standing on a street corner would hear its normal pitch

(A) doubled.
(B) halved.
(C) raised by about ten percent.
(D) lowered by about ten percent.
(E) lowered by about twenty percent.

42. You hear a distant police siren rise in pitch for a time, then remain at a lower pitch for a shorter time, then decrease in pitch for another period of time. The most likely explanation for the behavior is that

(A) the car is accelerating towards you, then stops, then is accelerating away.
(B) the car is driving around in a circle at a constant speed.
(C) the car is driving around in a circle, accelerating while it approaches you and decelerating as it recedes.
(D) the car is accelerating away from you, stops, then accelerates away.
(E) the car is moving toward you at a constant velocity, stops, then moves away at a constant velocity.

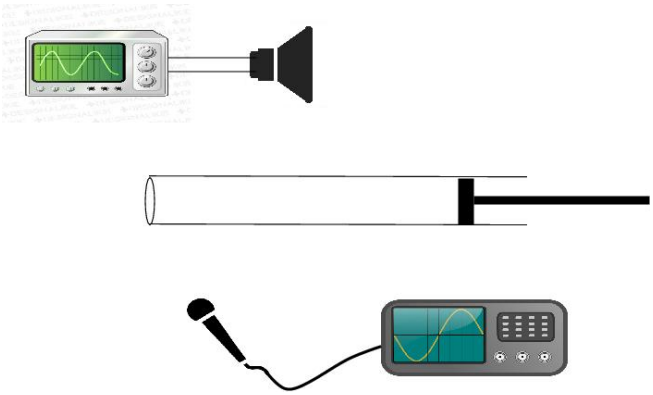
43. A police car is moving south on the expressway at $40\ \mathrm{m/s}$. You are moving south on the expressway in your own car at $30\ \mathrm{m/s}$. If the police siren emits a sound at 650 Hz, you will hear it at approximately

(A) 783 Hz
(B) 669 Hz
(C) 516 Hz
(D) 639 Hz
(E) 1032 Hz

Free response questions

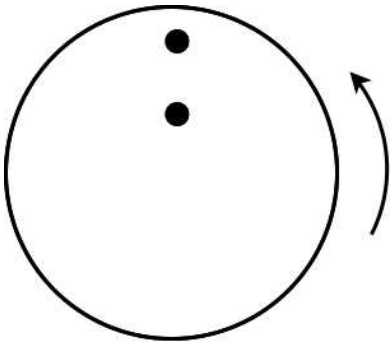
44. You are given two speakers that emit sound in phase and at the same frequency, in the vicinity of 1000 Hz. You are also provided a stopwatch, a tape measure, a meter stick, speaker stands and a sound-level meter. Describe an experiment you could do to measure the frequency of the speakers. Be sure you include a list of the equipment you have chosen and write down any equations you need. The speed of sound in the room can be taken to be 340 m/s.

45. As shown to the right, you are supplied a) a waveform generator that can generate a range of sinusoidal frequencies and that is attached to a speaker; b) a tube that is open at one end and closed at the other by a movable piston; c) an oscilloscope and microphone that can measure waveform amplitudes; d) a meter stick (not shown).



Design an experiment to determine the speed of sound in the laboratory.

46. Two identical coins are placed on a circular table at two radii r from the center of the table, as shown. As seen from the top, the table is rotating counterclockwise with a period T.



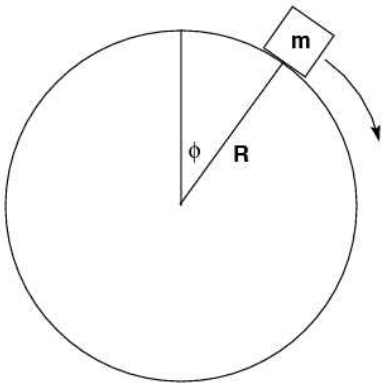
a) Draw in the vectors for the instantaneous linear acceleration of the outer coin, its instantaneous linear velocity and its angular velocity on the diagram.

b) The rate of rotation of the disk is gradually increased. If μ is the coefficient of static friction between the coins and the table, what is the condition for a coin to begin slipping in terms of μ, the period and any constants you might need. Which coin slips first?

c) If a third identical coin is glued on top of one of the others, how does the slip condition for that coin derived in part (b) change?

d) Assume that there is a coefficient of kinetic friction between one of the coins and the turntable. A short time after the coin begins to slip, in what direction is the force of friction and in what direction does the coin begin to move?

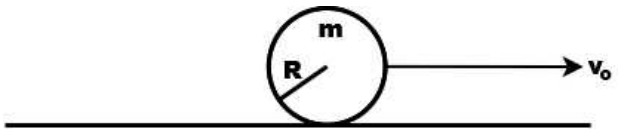
47. A block of mass m begins sliding off the top of a perfectly smooth sphere of radius R, as shown below.



a) At what angle ϕ does the block fall off the sphere?

b) If a coefficient of friction μ is introduced between the sphere and block, does the angle increase, decrease or stay the same?

48. A cue ball on a billiard table has mass m and radius R. You strike it with a cue stick, giving it an initial velocity v_o. At first it skids across the table, then begins to roll. The coefficient between ball and table is μ.



What is the velocity v_r in terms of v_o at which the ball begins to roll?

49. Light and radio waves exhibit the Doppler shift. Assume that radio waves obey the same equations for the Doppler shift as do sound waves. You are observing satellite signals from an orbiter in a circular orbit around a distant planet. You observe that the frequency of the signals rises and falls sinusoidally, reaching a maximum value of $32.000372\ \text{GHz}$ every 5 hours. You know that the transmission frequency of the satellite is $32.000000\ \text{GHz}$. (You may assume that the satellite is orbiting the planet edge-on with respect to your line of sight.)

a) What is the orbital velocity of the satellite?

b) What is the orbital radius of the satellite?

c) What is the mass of the planet?

6. Reference Frames, Relative Motion and Relativity Review

The subject test poses basic questions about special relativity and accelerated frames, which are covered in Sections 4 and 5 of this chapter. The remaining sections are at the AP level.

1. Essential Terms

In the previous chapters we have tacitly assumed that when describing displacement, velocity and acceleration, it is obvious what we mean if we say that an object "has a velocity of five meters per second." It is not obvious. All motion must be measured with respect to some **frame of reference,** or point of view. The motion of an object may appear very different when viewed from the ground and when viewed from a train moving with respect to the ground. Quantities such as kinetic energy and momentum assume different values in different reference frames. **All motion is relative to a chosen frame of reference**. Albert Einstein's **special theory of relativity** is based on this concept: motion must be described with respect to one reference frame or another. Conceptually, relative motion is a difficult subject because it requires you to visualize a system simultaneously from different reference frames. Practically, physics problems are frequently much easier to solve in one reference frame than another, and it is important to learn how to switch points of view, or **transform**, between frames. AP exams do pose questions that expect you to exploit such opportunities. The SAT subject test expects you to know a few basic facts about the theory of relativity and about the **principle of equivalence**, which we shall present in Sections 4 and 5.

2. Reference Frames and Relative Velocity

A **frame of reference** may be loosely thought of as a point of view. In physics, it is best thought of as a literal frame—a set of xy-coordinate axes in which an observer (yourself) stands at the origin. A reference frame may be attached to the ground, or to a train. The essential point is, it is always attached to something, and that something must be specified.

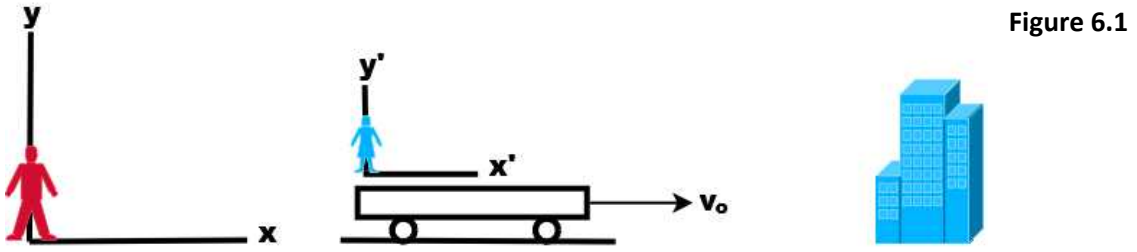


Figure 6.1

In Figure 6.1, the building appears at a constant distance to Bob in the xy-frame (usually called the "unprimed" frame), which is attached to the ground. A train is moving to the right at velocity v_o with respect to the ground. To Alice on the train, in the $x'y'$-frame (the "primed" frame), the building appears to be moving toward her at a velocity $-v_o$. Thus, the motion of the building appears different in the two frames.

As a second illustration, suppose that at the moment the train passes him, Bob tosses a ball to the right with a speed v at an angle θ from the horizontal such that the ball's x-velocity, $v\cos\theta$, equals v_o, the velocity of the train. Bob sees the ball follow a parabolic trajectory, but since the ball keeps pace with the train, Alice sees it move only up and down.

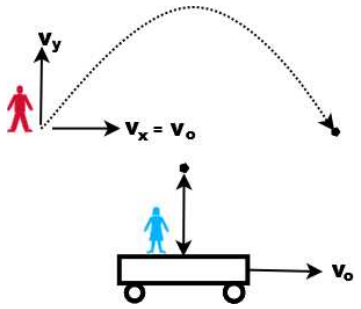


Figure 6.2

Another way of saying the same thing is that the relative x-velocity between the ball and the train is zero. The **relative velocity** between two objects a and b is simply defined as the difference between their velocities:

$$\mathbf{v}_{rel} \equiv \mathbf{v}_a - \mathbf{v}_b \quad (6.1)$$

The relative speed is simply the magnitude of the relative velocity: $v_{rel} = |\mathbf{v}_a - \mathbf{v}_b|$. Relative speed is often more convenient than relative velocity, since you aren't concerned with direction or which of the two objects is moving faster.

You know from daily experience that if you are in a subway car beneath Manhattan and another train passes by, then the passengers in the other train are hardly moving. The same is true when you are driving at 100 km/hr on a highway and another car passes you at 120 km/hr. Both these speeds are measured with respect to the ground, but the speed of the passing car you see from your own car is the relative speed of 20 km/hr. On the other hand, if the car is coming at you in the opposite direction, the relative speed is 220 km/hr, which is why head-on collisions often prove fatal.

Example 6.1. Assume that on a windless day, raindrops fall straight down with a speed of 9 m/s. On such a day, you hop in your car and drive north with a speed v_c, carefully observing that the drops fall at an angle $\phi = 40°$ from the vertical. How fast are you driving? What is the total speed of the rain as measured from the car?

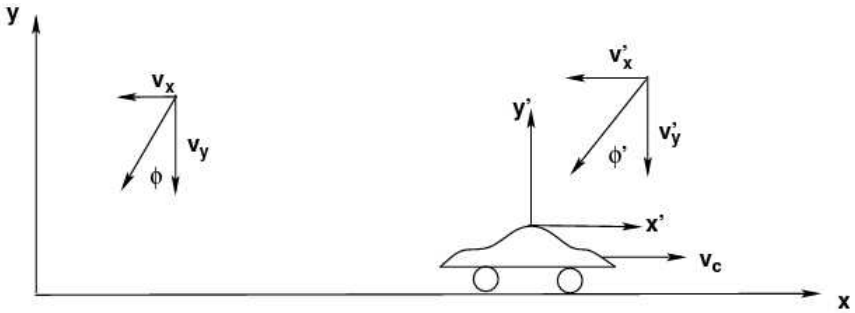


Figure 6.3

Figure 6.3 illustrates the problem. Call the frame of the ground the unprimed (xy)-frame. In this frame the rain falls straight down, making an angle $\phi = 0$ from the vertical. The rain has no x-velocity. Call the car's frame the $x'y'$-frame. In the car's frame, moving north, the rain appears to have an x-velocity equal to the car's speed v_c, but southward. From the figure, $\tan\phi' = v'_x/v'_y$. We are given that $\phi' = 40°$ and that $v'_y = 9$ m/s. Thus, the speed of the car is $v_c = v'_x = 9\tan 40° = 7.55$ m/s. Note that, consistent with the above discussion, the relative speed of the car and rain is just the speed of the car, since the x-velocity of the rain with respect to the ground is zero. The total speed of the rain measured in the car's frame is $v = \sqrt{v'^2_x + v'^2_y} = \sqrt{7.55^2 + 9^2} = 11.7$ m/s.

The phenomenon that rain falls at different angles in different reference frames is termed **aberration**. You have observed aberration yourself: on a wind-free day you hold an umbrella straight up, but if you run through the rain, or if the wind is blowing, you must tilt the umbrella to catch the raindrops.

Example 6.2. While driving north at 25 m/s during a windy rainstorm you notice that the rain makes an angle of 80° from the vertical. You turn around and drive south at the same speed. On this southbound leg you see that the rain falls straight down. What is the speed and angle of the raindrops relative to the ground? Assume that on a windless day, drops fall vertically.

Figure 6.3 also illustrates this problem. As shown on the left, in the ground's (xy)-frame, the velocity of a raindrop has x- and y-components, v_x and v_y and makes an angle ϕ with the vertical. The car's ($x'y'$-frame) is moving north at a speed v_c. In the car's frame, a raindrop has velocity components v_x' and v_y' and makes an angle ϕ' with the vertical. As in Example 6.1, $\tan\phi' = v_x'/v_y'$ and $\tan\phi = v_x/v_y$, but because the car is moving only in the x-direction, the raindrop's y-velocity is the same in both frames: $v_y' = v_y$. Furthermore, because from the car while going south the rain is seen to fall straight down, the car's velocity must equal the x-velocity of a raindrop, 25 m/s, south. (This is exactly the same situation as Bob tossing a ball with the x-velocity of the train, above.) Thus, when the car is moving north, the relative velocity of the car and the rain is $v_x' = 25 - (-25) = 50$ m/s. Then $\tan\phi' = \tan 80 = 50/v_y'$, or $v_y' = 8.8$ m/s. We already know that $v_y = v_y'$, which implies that $\tan\phi = 25/8.8$. Taking the arctan gives $\phi = 70.6°$, the angle as seen from the ground. The speed of the drops relative to the ground is $v = \sqrt{v_x^2 + v_y^2} = 26.5$ m/s.

Example 6.3. Alice is traveling due east at $v_a = 15$ m/s in car a. Bob is traveling due north in car b at $v_b = 20$ m/s. Bob tosses a ball perpendicularly out his window (as he sees it) with a speed of $v_t = 10$ m/s. At what velocity (speed and angle) is the ball launched from the frame of reference of the ground? At what velocity is the ball launched from Alice's frame?

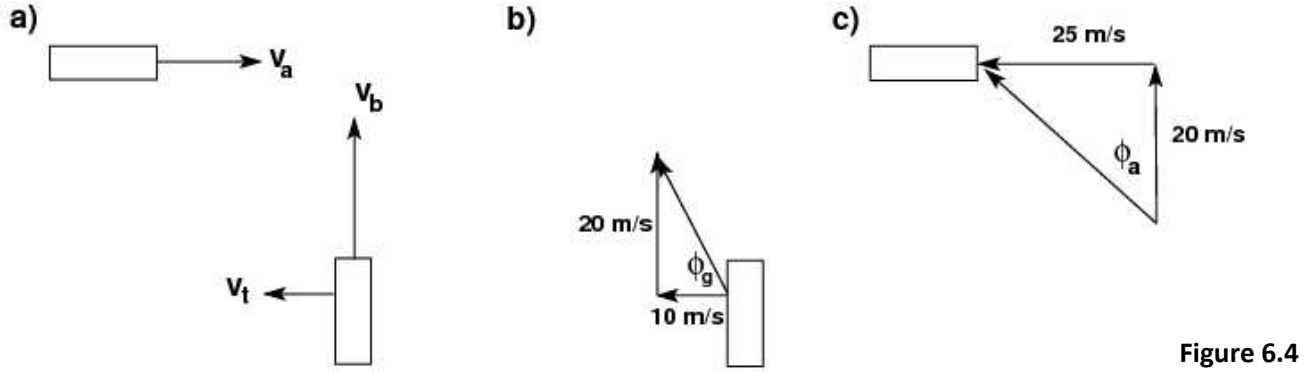


Figure 6.4

Figure 6.4a illustrates the general situation. From Bob's frame of reference, he throws a ball perpendicularly out the window with a speed $v_t = 10$ m/s. Someone observing from the frame of reference of the ground sees the ball moving west with a speed of 10 m/s and north with a speed of 20 m/s, as shown in Figure 6.4b. Thus, this observer measures the launch angle to be such that $\tan\phi_g = 20/10 = 2$. Taking the arctan gives $\phi_g = 63.4°$. The total speed measured from the ground is $v_{tg} = \sqrt{10^2 + 20^2} = 22.4$ m/s.

Alice sees the ball launched with a relative velocity of $15 - (-10) = 25$ m/s, west, and 20 m/s north. Thus, the speed she measures is $v_{ta} = \sqrt{20^2 + 25^2} = 32$ m/s. The launch angle from her frame is given by $\tan\phi_a = 25/20$, or $\phi_a = 51.3°$. We see that the answers vary considerably from one frame to another.

3. Transforming Between Frames

The above discussion can be made more formal to give a prescription for transforming from one frame to another. Returning to Figure 6.1, imagine that a bird has a displacement vector x in Bob's frame of reference and a displacement vector x' in Alice's frame. (For simplicity, we do this in one dimension and can dispense with vector notation.) The train itself is at position x_o in Bob's frame. Then, as can be seen from Figure 6.5, below,

$$x' = x - x_o \quad (6.2)$$

which relates the position of the bird in Alice's frame to the position of the bird in Bob's frame.

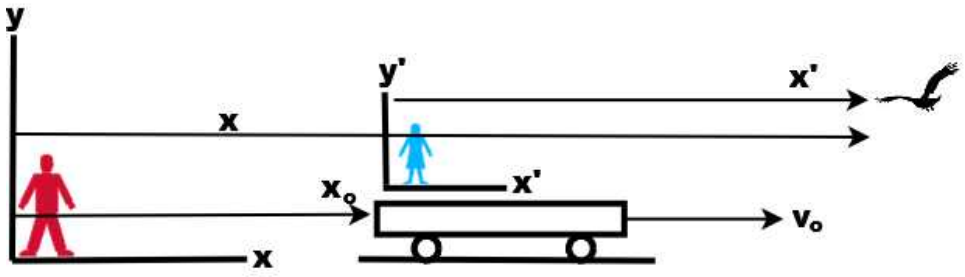


Figure 6.5

If the train is moving at a constant velocity v_o with respect to the ground, then the position Bob measures for the edge of the train is just $x_o = v_o t$, assuming that the train started out at the origin of Bob's frame. Substituting into the above equation gives,

$$x' = x - v_o t \quad (6.3)$$

Because velocity is the change of displacement with respect to time, the velocity of the bird measured in both frames is:

$$v' = v - v_o \quad (6.4)$$

Notice that the velocity of the bird as measured from the train (v') is just the relative velocity of the bird and train as measured from the ground, as stated above. Because acceleration is the change of velocity with time, if we now ask for the acceleration of the bird as viewed from both frames we get

$$a' = a \quad (6.5)$$

since we have assumed that v_o is constant and so the train's acceleration is zero. If the bird is moving in three dimensions, one merely has to write Eqs. (6.3)-(6.5) in vector notation: $\mathbf{x}' = \mathbf{x} - \mathbf{v}_o t$; $\mathbf{v}' = \mathbf{v} - \mathbf{v}_o$; $\mathbf{a}' = \mathbf{a}$ and work out each component separately, but in basic freshman problems, one rarely encounters such situations.

As simple as it looks, Eq. (6.5) is a profound result. It says that in two frames moving at a constant velocity with respect to one another, the acceleration measures in both frames are the same. Because $\mathbf{F} = m\mathbf{a}$, this in turn means that **the forces measured in two frames moving at a constant velocity with respect to one another are identical: $\mathbf{F}' = \mathbf{F}$.**

We have described a very special situation. If the train began to accelerate with a constant acceleration a_o, Eq. (6.3) would become $x' = x - v_o t - 1/2\, a_o t^2$. Eq. (6.4) would become $v' = v - v_o - a_o t$, and Eq. (6.5) would become $a' = a - a_o$. In other words, the acceleration the two observers assigned to the bird would differ in the two frames and, consequently, the forces measured by the observers in the two frames would also differ. *Newton's laws in the simple form* $\boldsymbol{F = ma}$ *would not hold*.

For this reason, essentially all problems in elementary physics and on exams are confined to reference frames moving at constant velocity relative to one another. Such special frames are termed **inertial reference frames** and the transformations (6.3-6.5) are termed **Galilean coordinate transformations**.

4. Special Relativity

The SAT subject test expects you to know a few basic results from Einstein's **special theory of relativity**. To formally derive these results is above the level of this book. However, we can give a brief idea of their origin.

Like the above examples, *special relativity deals with reference frames moving at constant velocity with respect to one another*, *in other words, with inertial frames*. However, Eq. (6.3) contains an unwarranted assumption. In writing it we, along with Newton, have tacitly assumed that the time t is measured to be the same in both frames, in other words that there is some universal, absolute time agreed upon by all observers. In the first years of the twentieth century Henri Poincaré and Albert Einstein realized that there is no reason to assume that this is the case.

Furthermore, if light acted like any other particle, we would expect its velocity, usually called c, to change depending on the reference frame, just as the bird's velocity in Section 3 changed depending on whether it was viewed from the frame of the ground or the train's frame. However, based on the fact that in **Maxwell's theory of electrodynamics** (Section 11.10) the speed of light appeared to be constant and that none of the experiments carried out in the nineteenth century were able to detect any change in the speed of light between reference frames, Einstein declared two postulates:

1. The laws of physics must be the same in all inertial reference frames.

2. The speed of light is measured to be the same constant, 3×10^8 m/s in *all* inertial reference frames.

The first postulate is termed the **principle of relativity** and it states that there is no preferred, or absolute, reference frame. The principle of relativity is entirely consistent with the discussion in the previous section, where we concluded that Newton's laws held in the same form in every inertial frame. The second postulate was revolutionary and had several far-reaching consequences:

1. There is no universal time. Observers in different inertial frames will measure events at different spatial coordinates (xy) as well as at different times *t*.

2. The lack of universal time implies that the Galilean transformations in Section 3 are *incorrect*. They work well at velocities small compared to the speed of light, but as velocities approach that of light they must be increasingly modified. At higher velocities one must employ a type of transformation between frames known as the **Lorentz transformations** (after Hendrik Lorentz, who discovered them before Einstein, but who had not based them on the two postulates.) The Lorentz transformations correctly describe the change of spatial coordinates, as well as the change in time intervals between events in different inertial frames.

3. No massive object's velocity can exceed, or even reach, the speed of light in vacuum, $c = 3 \times 10^8$ m/s. This statement implies, for example, that the Newtonian expression for kinetic energy, $KE = 1/2\, mv^2$, must be incorrect as well, because in this expression v can increase without limit and the kinetic energy can grow indefinitely.

4. Objects moving with respect to a stationary observer appear to shrink in the direction of motion. The formula is

$$L = L_o\sqrt{1 - v^2/c^2} \quad (6.6)$$

Here, L_o is what is called the "proper," or "rest-length," of the object: it is merely the length of the object in its own rest frame, that is, the length measured by someone to whom the object appears at rest. If the object is moving at velocity v with respect to another observer, that observer will measure the length to be the shorter value L. Notice that as v gets nearer the speed of light, L decreases towards zero. This famous effect is termed the **Lorentz contraction**.

5. Any observer in a clock's rest frame measures the time between clicks (1 second) to be $\Delta\tau$. This is termed "proper time." For example, if the clock is a wristwatch fastened to the observer's wrist, then he or she measures proper time. If another, identical clock, moves with respect to the observer at velocity v, then according to his or her own clock, the time between clicks of the moving clock is given by

$$\Delta t = \frac{\Delta\tau}{\sqrt{1-\frac{v^2}{c^2}}} \quad (6.7)$$

Notice that Δt is greater than $\Delta\tau$, which implies that according to the observer's watch, the moving clock is ticking more slowly. This famous effect is termed **time dilation**. We give an illustration of where this formula came from below.

6. $E = mc^2$. The mass of an object is equivalent to its energy content. The factor c^2 converts between the conventional units of mass (kg) and energy (joules); however, the famous formula shows that mass and energy are essentially the same thing. (Light, by definition, travels one light-year per year. In that set of units, $c = 1$ and $E = m$.) In this formula

$$m = \frac{m_o}{\sqrt{1-\frac{v^2}{c^2}}} \quad (6.8)$$

Where m_o is the "rest mass" of the object, the mass measured by someone in its rest frame. As v approaches c, m becomes infinite, which is why no material body can travel at the speed of light.

Example 6.4: The "light clock. Figure 6.6, below, illustrates time dilation.

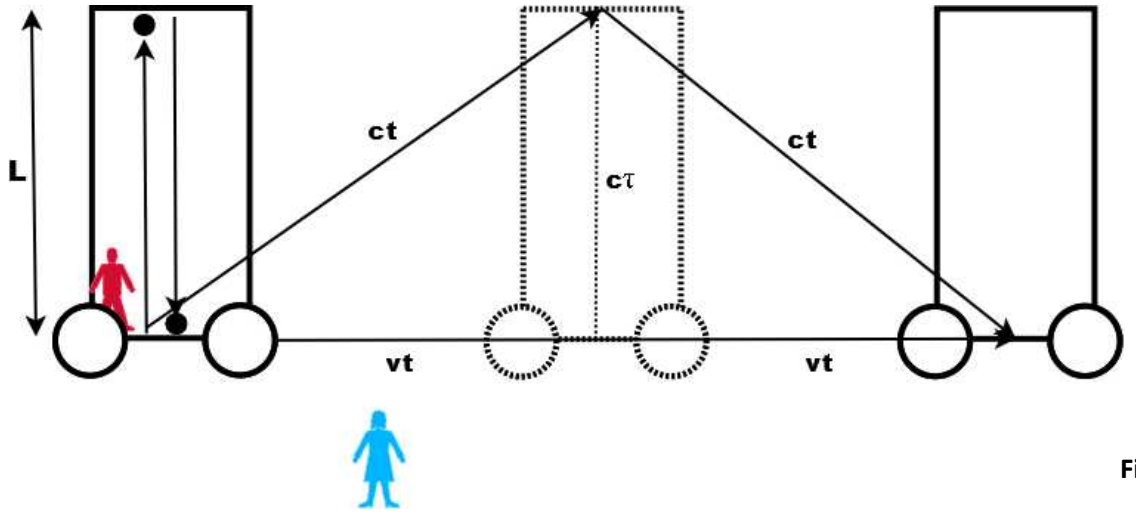


Figure 6.6

Suppose that a clock consists of two mirrors such that a ping-pong ball bounces between them, as on the left. The mirrors are separated by a distance L and the ball travels between them at a speed c, and so if we call τ the time it takes for the ball to travel between the mirrors as observed by Bob standing inside the clock, then $L = c\tau$. A roundtrip will take 2τ and we can call this roundtrip time one tick of the clock.

Suppose that the clock is mounted on wheels and moving to the right with a velocity v relative to the ground. As can be seen from the figure, Alice standing on the ground sees the ping-pong ball follow a diagonal trajectory, bouncing off the top mirror after the clock has moved a distance vt, where t is the time she measures by her own clock. If the speed of the ping-pong ball is always c, the distance it travels along the diagonal in this time is just ct, as shown. Since the hypotenuse of a right triangle is always the longest side, the time she measures between bounces of the ball must be longer than the time Bob measures. Indeed, by the Pythagorean theorem $c^2t^2 = v^2t^2 + c^2\tau^2$. Solving this equation for t gives

$$t = \frac{\tau}{\sqrt{1 - v^2/c^2}},$$

which agrees with Eq. (6.7), above.

This formula, however, was derived assuming that the ping-pong ball moves with speed c in both the clock's frame and the ground's frame. That is in fact the behavior of a beam of light according to Einstein's second postulate; it is constant regardless of the frame of reference.

Exercise: Show that if the ping-pong ball moves with velocity **c,** such that its y-component is c_o in both frames, but that its x-component is v in the ground frame (as would be the case with material objects obeying Newtonian physics), then $t = \tau$; there is no time dilation.

Exercise: Lay the light clock on its side and calculate the roundtrip time for the light-ball in the ground frame (Alice). Then use Eq. (6.7) for time dilation and derive Eq. (6.6), the Lorentz contraction. (This is fairly difficult but the solution to Problem 2.19 may help.)

Example 6.5: Cosmic-ray muons. Muons are subatomic particles that under normal circumstances decay in about 2.2 microseconds. According to Newtonian physics, even traveling at the speed of light, they would disappear after a distance $d = 3 \times 10^8 \times 2.2 \times 10^{-6} = 660$ m. However, muons are created in cosmic ray showers in the upper atmosphere. Since the 1940s it has been known that these muons can easily reach detectors at sea level, a distance of approximately 20 km from their point of creation. How is this possible?

Assume that the speed of the muons is $0.9997c$ relative to the ground. From the rest frame of the muons, the ground is rushing up at them at this speed. Thus, as far as the muons are concerned, according to Eq. (6.6) the distance between the point of their creation and the detector is Lorentz contracted to a distance $L = 20\sqrt{1 - .9997^2}$ km ≈ 490 m, which is less than 660 m, and so they can easily traverse it before decaying.

Alternatively, to a ground-based observer, the muon clock is moving and so the muon lifetime as measured by that observer is given by (6.7) as $\Delta t = \frac{2.2\times10^{-6}}{\sqrt{1-.9997^2}} \approx 9 \times 10^{-5}$ s. Traveling at essentially c, they are able to cross a distance of about 27 km before decaying, more than enough to reach the detectors at sea level.

Example 6.6: GPS. The Global Positioning System operates by triangulation. A GPS satellite sends out a radio signal that includes the exact time according to an onboard atomic clock. The radio signal, traveling at c, reaches the GPS unit in your phone, which records the precise time of arrival according to the phone's clock. If the satellite sends out a timing signal at exactly 12:00 noon and your unit receives it 6.67×10^{-2} s later, then using $d = ct$ the GPS system knows you are 20,000 km from the satellite. (**Exercise**: Verify this number.) Timing information from several satellites pinpoints your location.

However, at a GPS satellite distance of 26,400km from the center of the earth, Kepler's third law (Eq. 5.10) shows that the satellite has a velocity of about 3.9 km/s. Thus, $v/c = 1.29 \times 10^{-5}$, and $v^2/c^2 = 1.7 \times 10^{-10}$. Eq. (6.7) then tells us that the satellite clock and the phone clock are ticking at different rates such that $\Delta t \approx \Delta\tau(1 + 10^{-10})$, where Δt is a time interval (1 second) measured by the phone clock and $\Delta\tau$ is a time interval (1 second) measured by the satellite clock. In others words, according to special relativity, the satellite clock and the earthbound clock disagree by about 0.1 nanosecond each second. (**Exercise:** Verify all these numbers.)

This may seem like a tiny error, but the speed of light is large. In 0.1 nanosecond light travels 3 cm. Thus, since the GPS system calculates your position by $d = ct$, for each second ticked off by the phone, your position is in error by 3 cm. This error accumulates and so, if special relativistic time dilation were not taken into account, the GPS would be off by nearly two meters every minute, and within a few minutes become useless. (The situation is actually worse, since one also needs to correct for the gravitational field of the earth, which according to

Einstein's theory of gravitation, **general relativity**, also causes clocks to tick at different rates, by an amount twice the special relativity correction.)

5. The Principle of Equivalence

In Section 3.5, we mentioned that the fact that all objects fall at the same rate in a gravitational field is known as the **principle of equivalence**. If all bodies fall at the same rate, then you can jump onto a **freely-falling frame of reference**—a reference frame falling at the acceleration of gravity—and the bodies will no longer seem to be falling. They will be apparently weightless—a bathroom scale will be falling at the same rate as you and will not register your weight because you aren't pressing on it. The "weightlessness" of cosmonauts in orbit is precisely a state of **free fall**, in which all the bodies in the International Space Station are falling around the earth at the same rate.

If you were in an elevator and the cable snapped, you would momentarily be in a state of free fall and feel weightless. By the same token, you often feel heavier when an elevator accelerates upwards. To be sure, in a windowless elevator, you cannot tell whether you are accelerating upwards or whether the gravitational field of the earth has suddenly increased. In the case of the snapped cable, you cannot tell whether you are freely falling or whether the earth's gravitational field has suddenly disappeared. We are forced to conclude that locally **accelerations are equivalent to gravitational fields**. This is the principle of equivalence. The principle of equivalence is a law of nature. The special theory of relativity concerns objects moving at constant velocities; no accelerations are immediately involved. Einstein wanted to incorporate accelerations into this theory. Once he realized that accelerations were equivalent to gravitational fields, he understood that he was going to have to create a theory of gravity. Thus, the principle of equivalence laid the foundation for Einstein's general theory of relativity, the modern theory of gravitation.

Although the SAT subject test does not refer to the principle of equivalence by name, it does expect you to understand its basic consequences. The effective weight in a reference frame accelerating upward or downward is

$$W_{eff} = m(g - a) \quad (6.9)$$

where we have taken downward as positive. In other words, if you are in an elevator accelerating downward at $4.9\ \mathrm{m/s^2}$, your effective weight is one-half your normal weight. If you are in an elevator accelerating upwards at $9.8\ \mathrm{m/s^2}$, your weight is doubled.

6. Noninertial Reference Frames

Introductory physics courses avoid accelerating frames of reference—called **noninertial frames**—like the plague. They claim that Newton's laws do not hold in such frames and, as mentioned in Section 5.5, they dismiss common forces, such as the **Coriolis force** and the **centrifugal force** as figments of our imagination. Such forces arise only when we attempt to do physics in noninertial frames and they are therefore **fictitious**. But as we have just seen, gravity exists only when we do physics in certain frames; when we jump to a freely-falling frame, gravity disappears. By this argument gravity is also a fictitious force.

The fact is, one can do physics in any frame one chooses. Usually inertial frames make life easier and we choose them because in such frames Newton's second law takes the simple form $\mathbf{F} = m\mathbf{a}$. Nevertheless, one can write down the famous equation in a way that is valid in any reference frame, inertial or not, though we will not do it here.

In any case, the earth is a rotating frame of reference and the Coriolis force is responsible for hurricanes and so even if it is a figment of our imagination it is rather important. To see how this force arises, consider a space colony, as shown in Figure 6.7

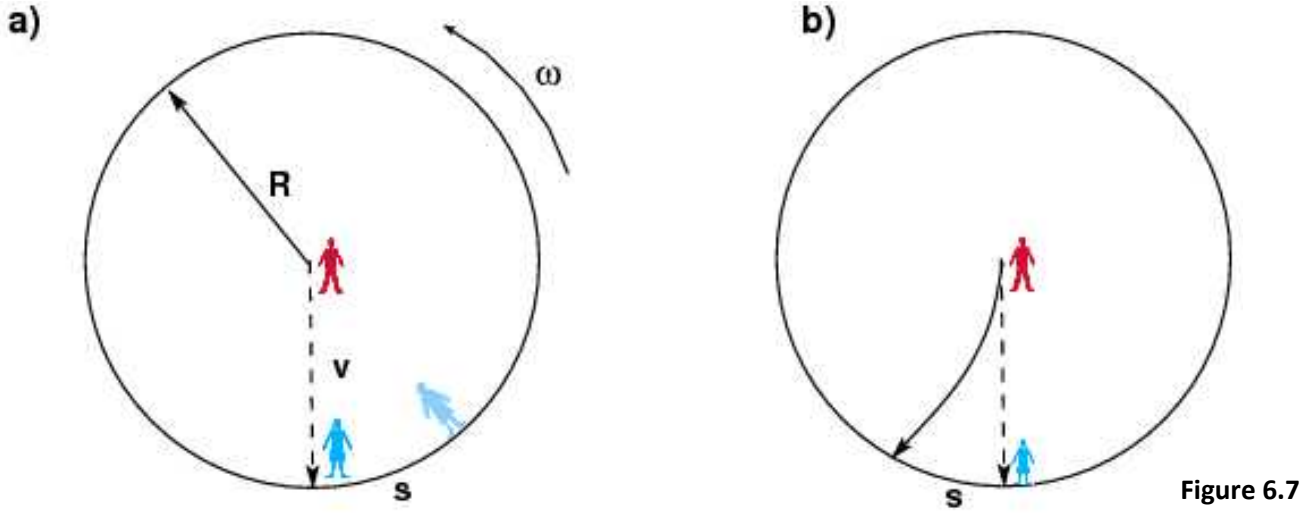


Figure 6.7

Imagine that you are viewing the colony from above, from the inertial frame of absolute space (Figure 6.7a). With respect to you, the colony is rotating with an angular velocity ω around a central axis. Bob, at the center, throws a ball with velocity v to Alice, standing beneath him at the rim of the colony. Once the ball leaves Bob's hand, no forces act on it (at least in your frame of reference) and you see the ball travel in a straight line (dashed) toward Alice. By the time it reaches the rim, however, she has rotated a distance s and cannot catch the ball.

As far as Alice is concerned however, she is standing happily and stationary on the rim of the space colony. Once the ball leaves Bob's hand, she sees it veer off to the left and land a distance s from her (Figure 6.7b). Being reluctant to believe that Bob is such a bad pitcher, she ascribes the deflection of the ball to a force. It is easy to calculate this force's magnitude. Assuming the acceleration is constant, we should have $s = 1/2\, at^2$, where t is the time the ball is in flight. If Bob throws with velocity v and the space colony's radius is R, then $t = R/v$. If v_t is Alice's tangential velocity at the edge of the colony, then from Eq. (4.2) we must also have $s = v_t t = R\omega t$. Equating the expressions for s and substituting R/v for t gives immediately $a = 2\omega v$. This is the **Coriolis acceleration** and $F = 2m\omega v$ is the **Coriolis force**, which is the force that deflects objects in a rotating frame of reference. The Coriolis force is essential in the formation of hurricanes and causes them to rotate counterclockwise in the northern hemisphere and clockwise in the southern hemisphere.

Example 6.7: Urban legend. A persistent urban legend is that the Coriolis force causes water to drain counterclockwise in sinks, bathtubs and toilets in the northern hemisphere and counterclockwise in the southern hemisphere. We can now see that this cannot be the case. The magnitude of the Coriolis acceleration is $2\omega v$. The relevant ω in this case is the angular rotation rate of the earth, 2π radians per day. From your number bank you remember that one day contains nearly 10^5 seconds. Therefore, $\omega \sim 6 \times 10^{-5}$ rad/s. The relevant v in this case is the velocity of the water in the bathtub, which might be 1 meter per second. Thus $a \sim 10^{-4}$ m/s^2. This is about 100,000 times less than the acceleration of gravity pulling the water into the drain. If you swish your hands back and forth about once a second over a distance of 30 cm, that also represents an acceleration of about 10 m/s^2 (**Exercise:** Convince yourself of this.) Thus, the Coriolis acceleration is utterly negligible to any other acceleration you will find in a bathtub. **Reality check:** Record which way the water drains in a bathtub for a month each time you take a bath.

If we imagine a system in which there are Coriolis accelerations, $2\omega v$, and other accelerations $a = v/t$, then the ratio of the accelerations is $a/a_c = v/2\omega vt$. This can be rewritten as $a/a_c = v/2\omega L$, where $L = vt$ is the characteristic size of the system. For a hurricane, L can easily be 500 km, or 5×10^5 m. The wind velocity of a

hurricane can be nearly 50 m/s. Thus, we find that $a/a_c \sim 1$, which tells us that Coriolis forces become equal to other forces on the scale of hurricanes, not bathtubs.

The other important force that appears in rotating frames of reference is the **centrifugal force**, from the Greek for "outwardly directed." Returning to the space colony, from your reference frame in absolute space, Alice is forced to follow a circular orbit only by the normal force that the wall of the colony exerts on her. If the wall disappeared, she would fly off in a straight line at her instantaneous tangential velocity. The normal force represents the centripetal force causing her to move in a circle.

On the other hand, from Alice's point of view, she feels the normal force of the wall pushing her toward the center of the colony. But she does not accelerate toward the center of the colony. Therefore, there must be an outwardly directed reaction preventing her from falling inward . It has exactly the same magnitude as the centripetal force, mv^2/R, but is pointed outward. Hence, centrifugal force. If Alice drops a ball, from your point of view above, it appears to move in a straight line with her tangential velocity. As far as she is concerned the ball falls to her feet until it hits the wall. Artificial gravity.

RELATIVE MOTION AND RELATIVITY PROBLEMS

Problems 1-11 concern the basic facts about relativity that one is expected to know for the subject test. The remaining questions are at the AP level but the topic of relative motion does not appear on the AP test.

1. An observer in a rocket watches another rocket pass by at near the speed of light. The observer will see passengers in the second rocket

   (A) retain their normal appearance.
   (B) shrink in the direction of motion.
   (C) shrink in the direction perpendicular to the motion.
   (D) shrink in both the direction of motion and the direction perpendicular to the motion.
   (E) expand in the direction of motion and shrink in the direction perpendicular to the motion.

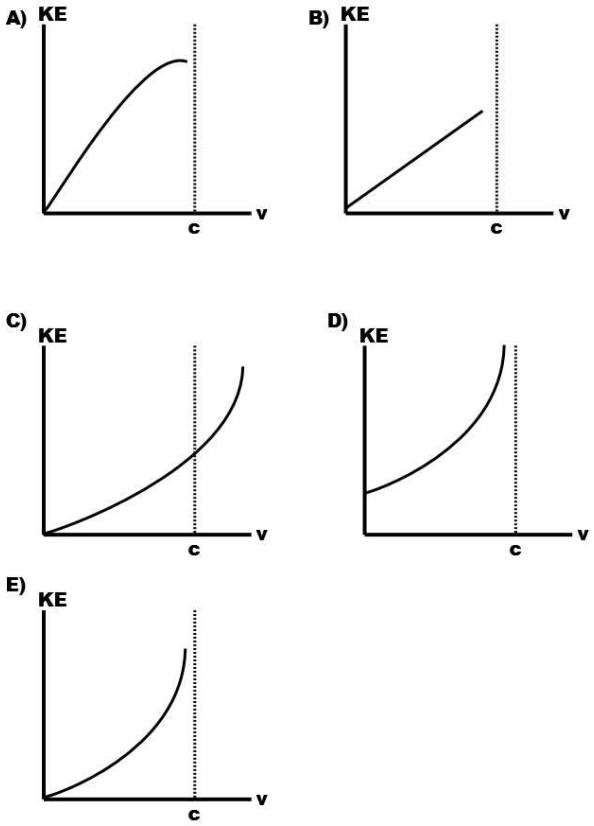
2. An observer in a space station watches a rocket fly past several times at various speeds that approach the speed of light. The rocket is carrying a clock on board. The observer in the space station compares the moving clock with her wristwatch and sees

   (A) that one second on the moving clock lasts longer than one second ticked off by her wristwatch.
   (B) that one second on the moving clock is shorter than one second ticked off by her wristwatch.
   (C) that a second on both watches lasts the same amount of time.
   (D) that the length of a second on the moving clock depends on how fast it is moving compared to the speed of light.
   (E) (A) and (D)

3. The Flash can run at $v = 6 \times 10^7$ m/s. To a resident of San Francisco, the length of the Golden Gate Bridge is about 2.7 kilometers. Its length to the Flash while running across it is closest to

   (A) 0.85 km
   (B) 2.45 km
   (C) 2.65 km
   (D) 3.0 km
   (E) 3.25 km

4. Which of the below graphs might represent the kinetic energy of a particle as a function of its speed? (In these graphs c represents the speed of light.)



(A) A
(B) B
(C) C
(D) D
(E) E

5. Which statements are true about the equation $E = mc^2$? (More than one response permitted.)

(A) The energy in matter is very large because the speed of light c is so large.
(B) The equation states that matter and energy are essentially the same thing.
(C) The entire mass of an object can be converted into energy.
(D) Energy can always be converted into mass.
(E) Conservation of energy and conservation of matter should be combined into one law.

6. The Empire State Building has a mass of approximately 3.2×10^8 kg. If the mass of a penny (2.5 gm) were entirely converted to energy, how high could you raise the building?

(A) About 7 km
(B) About 10 km
(C) About 25 km
(D) About 70 km
(E) About 100 km

7. If the Global Positioning System (GPS) did not take into account relativity,

(A) the GPS satellites would fall out of the sky.
(B) your GPS unit would always give the wrong position by about 10 meters.
(C) your GPS unit would always give the wrong position by about 50 meters.
(D) your GPS unit would give a position error that was always increasing.
(E) your GPS unit would always give the wrong position by about 100 meters.

8. Physical objects cannot travel faster than the speed of light c because

   (A) it is one of the postulates of special relativity.
   (B) it would take an infinite amount of energy to accelerate an object to c.
   (C) the length becomes infinite.
   (D) objects disintegrate before they reach the speed of light.
   (E) It is possible to travel faster than the speed of light.

9. If you are traveling north on a highway at 75 km per hour and another car is traveling south at 85 km per hour, then from your perspective, the southbound car is

   (A) moving at 160 km/h, southward.
   (B) moving at 10 km/h southward.
   (C) moving at 10 km/h northward.
   (D) moving at 160 km/h, northward.
   (E) None of the above.

10. You are standing on a horizontal platform that can accelerate up or down. If your mass is 60 kg and the platform is accelerating upward at $2.45\ \mathrm{m/s^2}$, your apparent weight is closest to

    (A) 470 N
    (B) 588 N
    (C) 735 N
    (D) 835 N
    (E) 1000 N

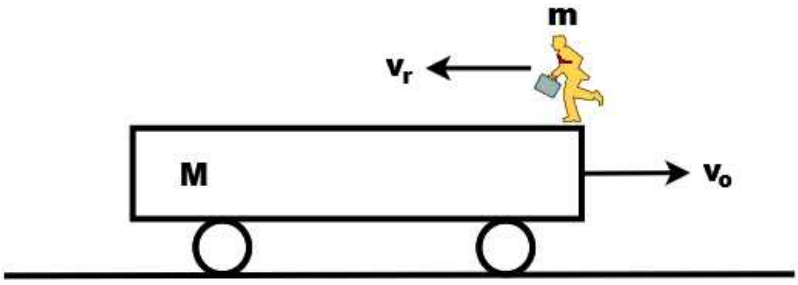
11. The same platform as in Problem 10 is accelerating downward at $2.45\ \mathrm{m/s^2}$. Your apparent weight is now closest to

    (A) 470 N
    (B) 588 N
    (C) 735 N
    (D) 835 N
    (E) 1000 N

Free response questions

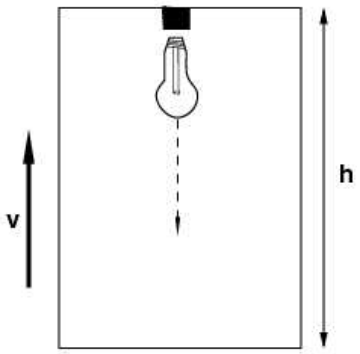
12. The Millennium Falcon drops out of hyperspace traveling at 100 km/s only to find a Battlecruiser directly ahead at a distance of 300 km. The Battlecruiser is moving at a constant velocity of 40 km/s in a straight line away from the Millennium Falcon. Han Solo immediately applies the brakes. What acceleration does the Falcon require to just avoid a collision with the Battlecruiser?

13. A man of mass $m = 70$ kg is standing motionless on a railway cart of mass $M = 1000$ kg, as shown.



Together they are moving with a velocity $\mathbf{v_o} = 4$ m/s to the right relative to the ground. At a time $t = 0$ the man begins running with a velocity $\mathbf{v}_r = 3$ m/s to the left, relative to the train. Assuming that there is no friction in the system, what is the final velocity of the train car $\mathbf{v}_c$ with respect to the ground?

14. As shown below, an elevator of height h moves upward at a constant velocity v. Vibrations cause the elevator's light bulb to fall from its fixture at time $t = 0$. The bulb hits the floor at time t later. Call the mass of the bulb m.

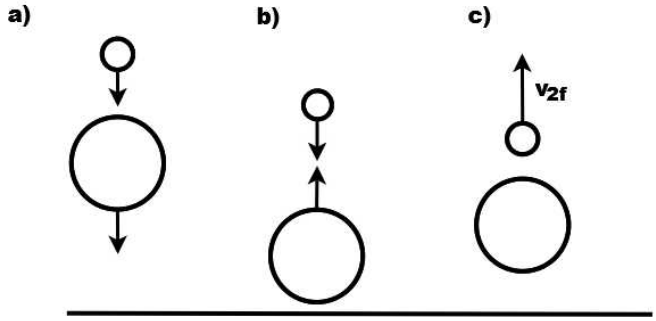


By direct calculation show that $\Delta KE = Fd$ holds in both the frame of reference of the elevator and the frame of reference of the ground. In this way verify that the work-energy theorem (Section 3.9) holds in both frames.

15. We showed in Problem 3.74 that during an one-dimensional elastic collision, the relative speed of the colliding objects is preserved. That is, $v_{1o} - v_{2o} = v_{2f} - v_{1f}$, where v_{1o} and v_{1f} are the initial and final velocities of the first object and v_{2o} and v_{2f} are the initial and final velocities of the second object. (Since this is a 1-D problem, we can dispense with vector notation; the conclusion also holds in three dimensions.) Show that the same result holds in the center of mass frame, the frame moving with velocity

$$v_{CM} = \frac{m_1 v_{1o} + m_2 v_{2o}}{m_1 + m_2}.$$

16. Here is a classic problem that has appeared on the AP physics exam: A small ball of mass m (a superball) is placed atop a large ball of mass $M \gg m$ (a basketball). They are dropped from approximately the same height h and fall with approximately the same velocity, as shown in (a) below. The basketball hits the floor slightly ahead of the superball and rebounds in an elastic collision off the floor, as in (b). The two balls collide in another elastic collision and the superball receives a final velocity v_{2f}, as in (c). What height does the superball reach in terms of h?



7. FLUID MECHANICS REVIEW

With the exception of the definitions of density and pressure in Eqs. (7.1) and (7.2), this chapter is entirely at the AP level.

1. Essential Terms

In elementary physics, **fluid** refers to a liquid or a gas, substances that under normal circumstances can and do flow. As with the other phenomena covered thus far, fluids obey Newtonian mechanics, and so one does not need to introduce new physical laws to describe their behavior. However, fluids consist of enormous numbers of particles and cannot be viewed in any manner as point objects, in the way a block of wood might be. Fluids must be described "in bulk" by the basic quantities **density**, **pressure** and **temperature**. These three quantities, along with Newton's second law and conservation of energy adapted to fluids, are enough to account for the basic behavior of fluids, including **buoyant forces** (**Archimedes' Principle**) and **Bernoulli's equation**.

2. Density

If an object has a mass M and volume V, its mass *density* ρ *("rho") is defined as the mass per unit volume*, or

$$\rho = \frac{M}{V} \quad (7.1)$$

This equation strictly holds only when the density is constant over the volume of the object.[1] A fluid in which the density is constant is termed **incompressible**. **Most elementary physics problems involving liquids assume that the fluid is incompressible.**

Note that the dimensions of density are $\mathrm{M/L^3}$. In the SI system, density is measured in $\mathrm{kg/m^3}$, which is inconvenient for talking about ordinary substances and one often encounters density expressed in $\mathrm{gm/cm^3}$. No one's name is attached to a unit of density.

Tip: For your number bank, it is useful to remember that the density of most ordinary materials lies between 1 and 10 grams per cubic centimeter, in other words between 1000 and 10,000 kg per cubic meter. This will provide a reality check when you calculate densities.

The density of water is 1 gm per cubic centimeter. Most metals have a density in the range of 8 $\mathrm{gm/cm^3}$ and in Example 0.4 we found that the average density of the earth is about 6 $\mathrm{gm/cm^3}$.

3. Pressure

If a force **F** is applied to a surface of area A, the *pressure is defined as the force per unit area*, or

$$p = \frac{F}{A} \quad (7.2)$$

[1] The precise definition of density is $\rho = dM/dV$, or equivalently, $M = \int \rho \, dV$, which must be used when ρ is not constant.

As with density, this definition is strictly true only when dealing with constant pressure.[2] Also notice that as defined, pressure is a scalar, so we are really talking about the magnitude of the force divided by the area. We regularly speak of "pressure exerted on a surface," although strictly speaking, since pressure is a scalar and has no direction, it is the force associated with the pressure that carries the direction. The dimensions of pressure are $\mathrm{M/LT^2}$ and the SI units are $\mathrm{N/m^2}$. The unit of pressure has received a name: Pascal, after Blaise Pascal: $1\ \mathrm{N/m^2} \equiv 1\ \mathrm{Pa}$.

Tip: The dimensions of pressure are force/area, which is the same as energy/volume. (See Problem 1.7.)

Example 7.1: What is the approximate pressure exerted by your feet on the floor? If your mass is 60 kg, then your weight is $W = mg$, or approximately 600 N. If the area of your foot is $8 \times 20\ \mathrm{cm^2} = 16 \times 10^{-3}\ \mathrm{m^2}$, then from Eq. (7.2), the pressure one foot exerts on the floor is $600 \div 16 \times 10^{-3} \approx 3.7 \times 10^4$ Pa. The total pressure from both feet is twice this value.

Example 7.2: What is the pressure at the bottom of a column of water of height h?

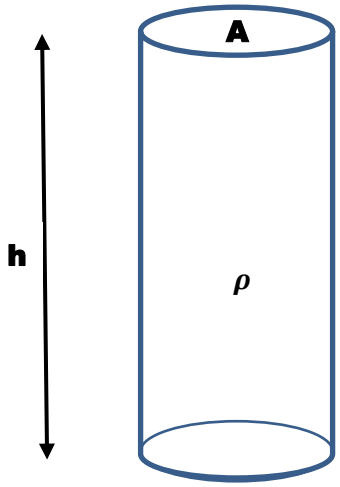


Figure 7.1

As shown in Figure 7.1, let the area of the column be A, the height be h, and assume that the density ρ is constant. The volume of the cylinder is $V = Ah$ and so the total mass is $m = \rho V = \rho Ah$. The force exerted by the column on a surface at the base is just its weight: $W = mg = \rho gAh$, and the pressure is by definition $F/A = \rho gh$. This is true of any fluid, giving the important expression for the pressure at the bottom of any fluid column with constant density:

$$p = \rho gh \quad (7.3)$$

Note that this is independent of the area, as it must be, since we divided out A. Thus, p is the pressure at the base of a column of fluid regardless of how wide it is, be it a soda straw or a swimming pool. *If an external force is acting on the fluid, then the pressure due to that force must be added to the pressure of the fluid*. For example, **if the fluid column is open to the atmosphere at the top, then at a depth h below the fluid surface $p = p_o + \rho gh$, for atmospheric pressure p_o.**

[2] The precise definition is $p = dF/dA$, or $F = \int p\, dA$.

Example 7.3: We can use Eq. (7.3) to estimate atmospheric pressure at sea level. A good number for your number bank is that the density of air is about 1 kg/m^3. Then, if we say that most of the air in the atmosphere is below an altitude of 10 km, Eq. (7.3) gives $p = 1 \times 10 \times 10^4 = 10^5$ Pa. This is actually quite close to the real answer: 101.3 kPa. (In fact the density of air depends on temperature and is nearly 1.3 kg/m^3 at the standard temperature of 0° C. It also varies with altitude.)

Pressure has several properties that become important in many problems:

1. The pressure at a given depth in a static fluid is the same in all directions.

(If this were not true then a small parcel of fluid would begin to accelerate in one direction or another, but this contradicts the assumption that the fluid is static.) For example, at a given depth, the fluid pressure acting up on an object is the same as the pressure acting down

2. Pressure in a static fluid acts perpendicularly to the walls of any vessel containing it.

(If the fluid exerted a force parallel to the walls, then by Newton III, the wall would exert an equal and opposite force on the fluid, and a fluid parcel would move parallel to the wall. This contradicts the assumption that the fluid is static.)

Important point 1, above, has many practical consequences, which are manifested in **Pascal's Principle**, although it is not a new physical concept.

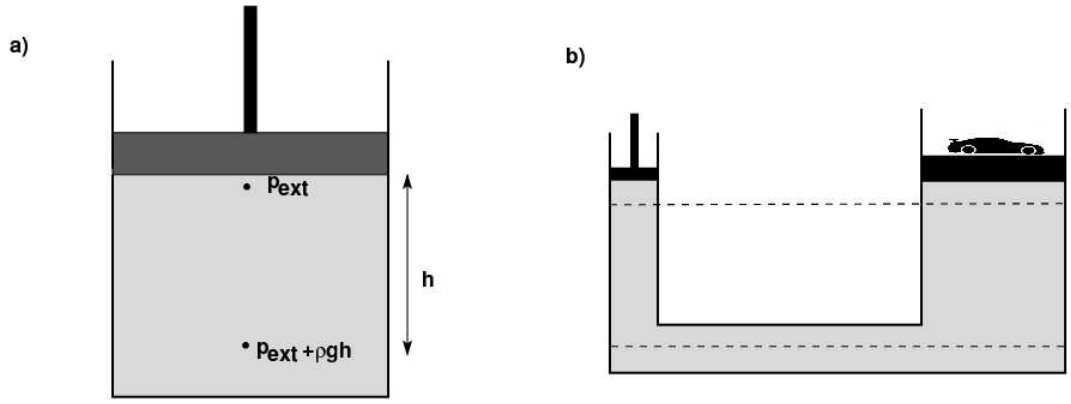


Figure 7 .2

Consider Figure 7.2a. A piston presses down on a fluid of density ρ in a container, exerting an external pressure p_{ext}. We know from above that the pressure at any height h is given by $p(h) = p_{ext} + \rho gh$. If the external pressure p_{ext} is increased—for instance by increasing the force exerted by the piston—then the change in pressure is $\Delta p(h) = \Delta p_{ext} + \Delta(\rho gh)$. However, in an incompressible fluid the density is by definition constant and so the last term is zero. The change in pressure at a depth h is therefore merely the change due to the external force, $\Delta p(h) = \Delta p_{ext}$. This statement is Pascal's Principle: **An external pressure applied to an enclosed fluid increases the pressure anywhere in the fluid by the same amount.**

Example 7.4: The previous result finds application in many aspects of hydraulics, in particular, the hydraulic jack, illustrated in Figure 7.2b. We know that the pressure anywhere along the dashed lines must be the same, since they are at the same height h in the fluid, and according to Important Point 1 above, the pressure is the same in any direction, including up. According to Pascal's Principle, if the piston increases the pressure on the left part of the upper dashed line, it is increased by the same amount on the right part. Thus, $p_1 = p_2$, where p_1 is the pressure under the piston and p_2 is the pressure under the car. Because pressure is force/area, $F_1/A_1 = F_2/A_2$, where A_1 is the area of the piston and A_2 is the area of the platform beneath the car. Then $F_2 = F_1 A_2/A_1$. If A_2 is much larger than A_1, we have created a force multiplier that is capable of lifting large weights, in particular automobiles. You have seen hydraulic jacks in action every time you have visited your local car repair shop.

4. Buoyant Force and Specific Gravity

From daily experience you know some objects float on water and that you feel lighter in water than on land. This is because the water is exerting a **buoyant force** on any object fully or partially immersed in it; the buoyant force pushes the object upward. We can easily calculate the buoyant force on an object as follows: Suppose the body is cylindrical with a cross-sectional area A, and a length y, as shown in Figure 7.3.

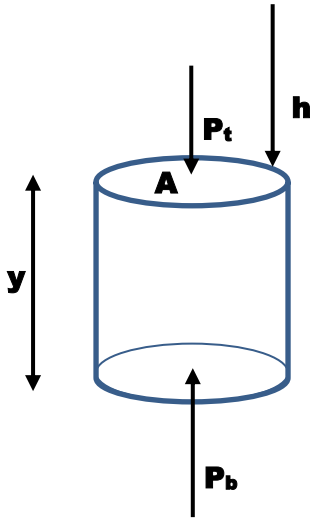


Figure 7.3

The top of the object is at a distance h below the surface of the fluid and the bottom is at a distance $h + y$. The fluid exerts a net force on the object $F = F_b - F_t$, where F_b is the force on the bottom acting upwards and F_t is the force on the top acting downward. Then $F = [p_b - p_t]A$, where p_b and p_t are the pressures at the bottom and top. (That there is a pressure acting upward follows from important property 1, above.) From Eq. (7.3), $F = \left[\rho_f g(h + y) - \rho_f gh)\right]A$, or $F = \rho_f gyA$, where we write ρ_f to indicate the density of the fluid.[3] But yA is just the volume V of the cylinder and $\rho_f V$ is merely the mass m_f of the fluid that would occupy V. The weight of the displaced fluid is $W_f = \rho_f Vg = m_f g$ and hence the net force acting upward on the object, the buoyant force, is

$$F_{buoyant} = \rho_f Vg = m_f g = W_f \quad (7.4)$$

We know that this force acts upward because the pressure at the bottom of the object is greater than the pressure at the top. The important equation (7.4) is known as **Archimedes' Principle** after the Greek mathematician who discovered it two thousand years ago, and it is usually stated as **the buoyant force acting on an object is equal to the weight of the displaced fluid.**

Although we have carried out this calculation for a cylindrical object, Archimedes' Principle is valid regardless of the shape of the immersed body, as can be seen by a simple argument. Suppose an irregularly shaped object is submerged in water, as shown on the left of Figure 7.4.

[3] If the fluid is exposed to the atmosphere, we should include the atmospheric pressure in this calculation, but since it gets added to both the top and bottom forces, it merely cancels out.

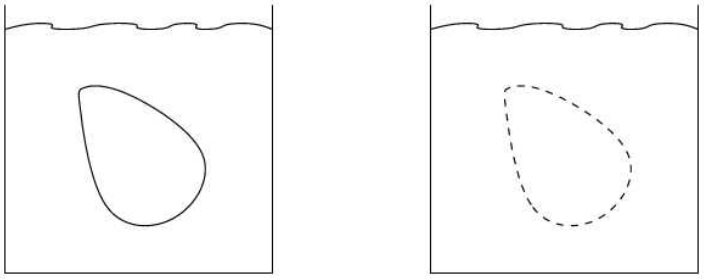


Figure 7.4

There will be some buoyant force F_b acting upwards on the body. We now replace the real object, which may be made of wood or metal, with an imaginary one of exactly the same size and shape, but made of water, as on the right. F_b hasn't changed since it is an external force and depends on only the external fluid. The imaginary object is in equilibrium with the water—since it is also made of water—and therefore the buoyant force on the imaginary object must be the object's weight: $F_b = m_f g$. But the amount of water in the imaginary object is exactly the same as the amount of water displaced by the real object; the buoyant force on the real object remains $F_b = W_f$, as stated.

If the object is only partially submerged to a distance y beneath the water, then the pressure at the top is p_o, atmospheric pressure, and the pressure at the bottom is $p_o + \rho_f g y$. Therefore, the buoyant force is $F_b = \rho_f g y A = \rho_f g V_s$, where V_s is the *submerged* volume. Hence, as before, the buoyant force is the weight of the displaced fluid. Consequently, **Archimedes' Principle applies whether the object is fully or only partially submerged.**

Due to the buoyant force, an object weighs less in water than in air. This fact gives rise to the concept of **specific gravity**. Consider an object of mass m, volume V and density ρ (no subscript). Its weight in air is $W = mg = \rho g V$. Immersed in water, however, two forces are acting on the body: the force of gravity downward and the buoyant force upward, as shown in Figure 7.5.

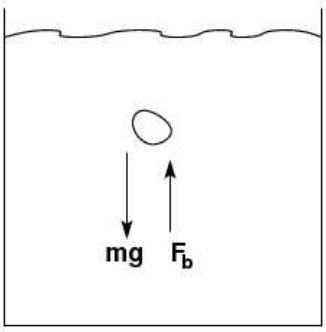


Figure 7.5

If the weight of the object in the water is W', then $W' = W - F_b$, giving $W - W' = F_b = \rho_f g V$. Since the weight in air $W = \rho g V$, dividing the latter equation by the former yields

$$\frac{W}{W - W'} = \frac{\rho}{\rho_f} \qquad (7.5)$$

The ratio $(\text{density object})/(\text{density water})$ is termed the **specific gravity** of an object, and Eq. (7.5) shows how to measure it: weigh the object in air and in water and form the ratio on the left-hand-side of the equation. Eq. (7.5) also shows that if an object has a specific gravity of 1, $W' = 0$, which implies the buoyant force on the object is equal to its weight in air; it "just floats," or has what is called neutral buoyancy.

Example 7.5: A 1500 kg statue of Archimedes is found in the water off Greece. If the tension in the cable used to raise it at constant velocity to the surface is found to be $T = 9256$ N, what is the specific gravity of the statue? What is it made from?

If the statue is being raised at constant velocity, the total force on it must be zero. Drawing in an upward arrow on Figure 7.5 to represent the tension in the cable, we see that the tension must equal the force of gravity minus the buoyant force: $T = mg - F_b$, or $F_b = mg - T$. But from the above discussion $F_b = W - W'$. With $W = mg$, Eq. (7.5) can then be rewritten as $mg/(mg - T) = s$, where $s \equiv \rho/\rho_f$ is the specific gravity. Plugging in the given values for m and T gives $s = 2.7$, which means the statue is almost certainly marble.

Example 7.6: What fraction of the volume of an iceberg is submerged under water? The density of seawater is approximately $1.03\ \mathrm{gm/cm^3}$. The density of sea ice is approximately $0.91\ \mathrm{gm/cm^3}$.

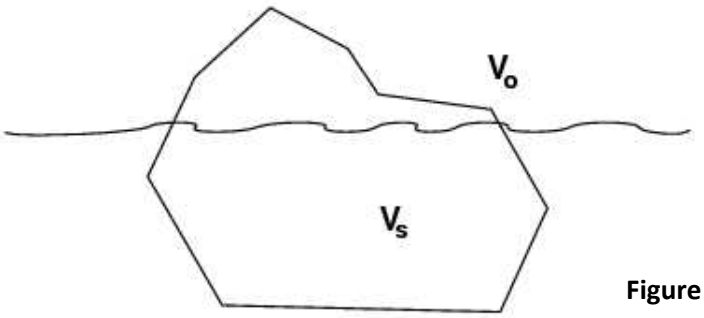


Figure 7.6

Let m_o and V_o equal the total mass and volume of the iceberg, and let V_s be the submerged volume, as shown in Figure 7.6. If the iceberg is floating in the water, then it is in equilibrium and the buoyant force must equal the iceberg's total weight: $F_b = m_o g = \rho_i V_o g$, where ρ_i is the density of the ice. But by Archimedes' Principle, the buoyant force is also equal to the weight of the displaced water, or $F_b = \rho_f V_s g$, where ρ_f is the density of water. Equating the two expressions for F_b gives

$$\frac{V_s}{V_o} = \frac{\rho_i}{\rho_f}.$$

In other words, the submerged fraction of an iceberg's volume is equal to its specific gravity. For the given numbers $V_s/V_o \approx 0.88$. Thus, about 90% of the iceberg is underwater, as the passengers on the Titanic unfortunately learned.

Example 7.7: An ice cube is floating in a glass of water such that the water level comes right up to the rim. As the cube melts, does the water overflow the glass?

This is a popular question that surfaces even on quiz shows. One thing on which everyone can agree is that by conservation of mass, the initial mass of the ice cube m_o must equal the final mass of the water m_f into which it melts. That implies $\rho_i V_o = \rho_f V_f$, where V_f is the final volume of the water. Hence, $V_f = V_o \rho_i/\rho_f$. But from the previous example $\rho_i/\rho_f = V_s/V_o$, giving $V_f = V_s$. In other words, after melting, the water occupies only the volume of the ice cube that was originally submerged. Therefore, the water level doesn't change.

Example 7.8: Displacement tonnage. The weight of naval vessels is often measured in "displacement tonnage," which is precisely the weight of the water that the ship displaces when its fuel tanks are full and stores are aboard. However, by Archimedes' Principle, this must also be the weight of the ship. Displacement tonnage should be distinguished from commercial "tonnage," which is the weight of the cargo on board.

5. The Equation of Continuity and Bernoulli's Equation

Suppose a liquid is flowing at velocity v through a pipe with cross sectional area A. In Section 3.14 we showed with slightly different notation that the number of particles flowing past A per second is $\Delta N/\Delta t = nAv$, where n is the number density of particles, or particles per cubic meter. Multiplying the above equation by the mass of a particle, m, gives the mass flow through the area: $\Delta m/\Delta t = \rho Av$.

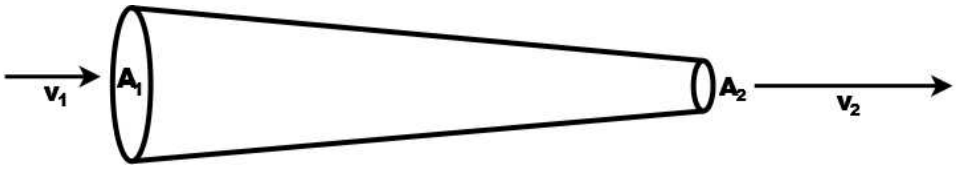


Figure 7.7

Now, mass cannot disappear while flowing through a pipe. If, as illustrated in Figure 7.7, a pipe has a changing area, such that the cross section is A_1 in one place and A_2 in another place, we must then have $\rho_1 A_1 v_1 = \rho_2 A_2 v_2$. If, furthermore, the fluid is incompressible, then the density is everywhere constant and

$$A_1 v_1 = A_2 v_2 \quad (7.6)$$

This important equation is known as the **equation of continuity** and is a statement of the conservation of mass in a fluid flowing through a pipe.

Example 7.9: Garden hose. You have undoubtedly covered one end of a garden hose with your thumb in order to make the water shoot farther. According to the equation of continuity, if you cover half the area of the end, then the water must double its velocity. As the time it takes for the water to hit the ground has not changed, it should shoot twice as far. An adjustable nozzle can close the end even more, further increasing the water velocity.

Closely related to the equation of continuity is **Bernoulli's equation**, after Daniel Bernoulli, which finds applications in a wide variety of situations. Bernoulli's equation explains why pitchers can throw curve balls and helps explain why airplane wings have lift. We will not formally derive the equation but one doesn't need to if one remembers that **Bernoulli's equation is an expression of conservation of energy for incompressible fluids.**

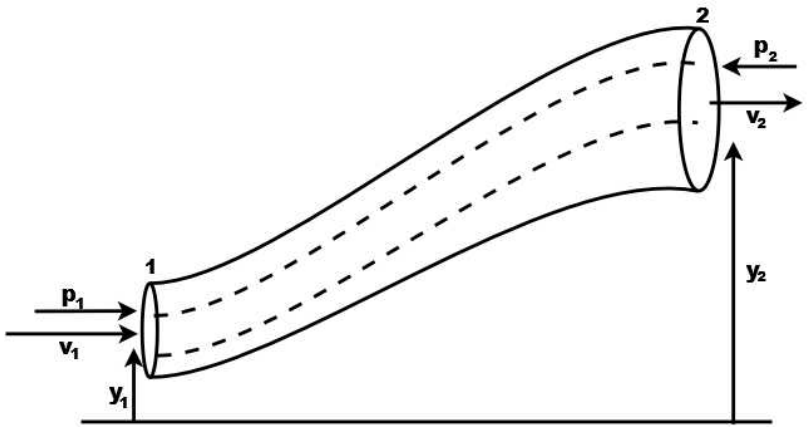


Figure 7.8

Figure 7.8 illustrates the setting in which Bernoulli's equation applies. The dashed lines represent **streamlines**, which are trajectories that follow the path of the fluid particles. *Bernoulli's equation compares the pressure and velocity of an incompressible fluid at any two points along a streamline*, for instance at point 1 on the left end of one of the illustrated streamlines, and at point 2 at the right end. Strictly speaking, Bernoulli's equation does *not* apply when comparing points on different streamlines. It is also assumed that the fluid is flowing smoothly, without turbulence, and any viscosity (friction) between the pipe and fluid is ignored.

In Section 3.9 we proved a simple version of the work-energy theorem. If you exert an external force F_{ext} on a mass m and displace it an amount s, then the work-energy theorem tells us that the work performed by F_{ext} must be the change in kinetic energy: $F_{ext}s = 1/2\, mv_2^2 - 1/2\, mv_1^2$. If some of the work goes into raising the mass against gravity from a height y_1 to y_2, then the net work is the work done by the external force minus the work done by the force of gravity: $F_{ext}s - mg(y_2 - y_1) = 1/2\, mv_2^2 - 1/2\, mv_1^2$. Rearranging gives $F_{ext}s + mgy_1 + 1/2\, mv_1^2 = mgy_2 + 1/2\, mv_2^2$. If $F_{ext} = F_1 - F_2$, where F_1 and F_2 are two forces acting on the mass in opposite directions we can further write

$$F_1 s + mgy_1 + 1/2\, mv_1^2 = F_2 s + mgy_2 + 1/2\, mv_2^2 \ .$$

For fluids the situation is nearly the same, except that a fluid is an extended body. There will be a fluid element to the left of the one illustrated in Figure 7.8 exerting a force to the right on the mass of fluid m_1 at position 1 and displacing it an amount s_1. Similarly, a fluid element to the right of the one illustrated will exert a force to the left on the mass of fluid m_2 at position 2 and displace it by an amount s_2. Thus, the above becomes

$$F_1 s_1 + m_1 g y_1 + 1/2\, m_1 v_1^2 = F_2 s_2 + m_2 g y_2 + 1/2\, m_2 v_2^2 \quad (7.7)$$

Now, each term in this equation represents an energy. If we divide the mass m of a fluid element by its volume, we get its density ρ. Furthermore, recalling the tip in Section 3, the dimensions of pressure are energy/volume. Therefore, if we divide the previous equation by the fluid volume, it is plausible that the above becomes

$$p_1 + \rho g y_1 + \frac{1}{2}\rho v_1^2 = p_2 + \rho g y_2 + \frac{1}{2}\rho v_2^2 \quad (7.8)$$

for any two points 1 and 2 along a streamline. This is Bernoulli's equation (or principle). Since points 1 and 2 are arbitrary, it can be restated as

$$p + \rho g y + \frac{1}{2}\rho v^2 = constant$$

along a fluid streamline. For points at a constant height $y = 0$, the result is often remembered as **pressure high, velocity low**. If the pressure is high at a point along a streamline, the velocity there must be low compared to another point.

Exercise: Assume that the fluid is incompressible and that $m_1 = \rho A_1 s_1$, where m_1 is the mass of the bit of fluid in position 1, A_1 is the area of the tube at that point, and s_1 is the distance that m_1 is displaced. With a similar expression for m_2, use the equation of continuity Eq. (7.6) to show that Eq. (7.7) results in Eq. (7.8), Bernoulli's equation.

Example 7.10: A large water tank of height h is open at the top to the atmosphere. If a small faucet is placed at the bottom of the tank, with what velocity will the water come out when the faucet is opened?

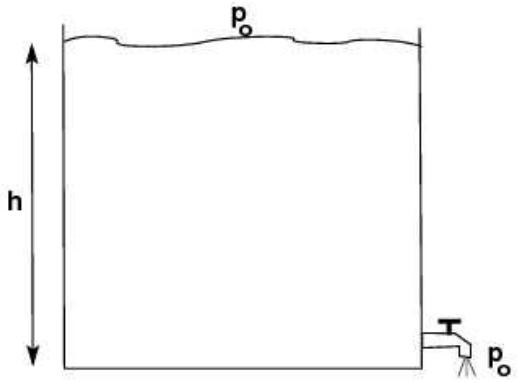


Figure 7.9

Let a point at the top of the tank be point 1 along a streamline and a point at the faucet be point 2 along the same streamline. When the faucet is opened, both points are at atmospheric pressure, p_o. Since the faucet is assumed to be small, the tank will not be draining quickly and we can take the velocity at point 1 to be approximately zero. Then Bernoulli's equation gives $\rho g(y_1 - y_2) \approx 1/2\rho v_2^2$, or $v_2 \approx \sqrt{2gh}$. Notice that this is just the velocity of an object dropped from a height h as it hits the ground, as one would expect from conservation of energy. This particular result is known as **Torricelli's Law**.

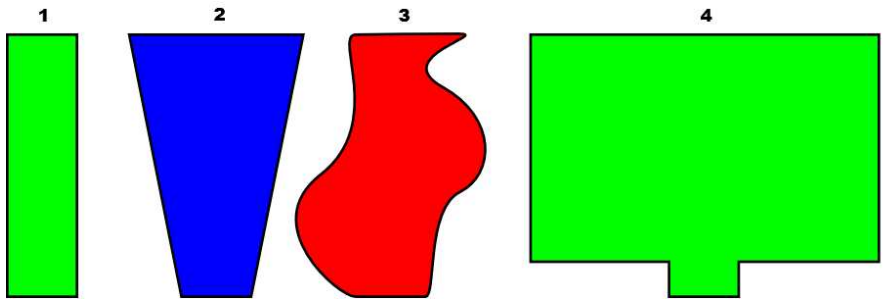
Example 7.11: Water in the basement utility room is fed into a pipe of radius 2 cm at a velocity $v = 1$ m/s and a pressure of 5 atmospheres. What is the water velocity and pressure in a pipe on the second floor when the water is flowing? The second-floor pipe has a radius 1 cm and is located 6 meters above the utility room.

Call the utility room point 1 and the pipe on the second floor point 2. According to the equation of continuity, $A_1 v_1 = A_2 v_2$. Because $A_1 = 4A_2$, we have $v_2 = 4v_1 = 4$ m/s. Substituting this into Bernoulli's equation gives $p_1 + \rho g(y_1 - y_2) + \frac{1}{2}\rho(v_1^2 - 16v_1^2) = p_2$. Recalling that atmospheric pressure is close to 10^5 Pa, and using $g = 10$ m/s^2 yields $p_2 \approx 4.3$ atmospheres.

FLUID MECHANICS PROBLEMS

Because this chapter is not covered by the SAT subject test, these are all AP-level problems.

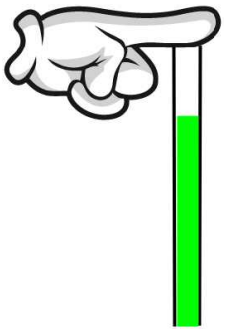
1. The three containers below all have the same height h and the same base area A.



If the containers are filled with water to h, the force on the base of the containers is, from highest to lowest:

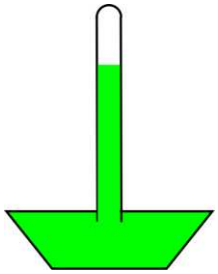
(A) 1,2,3,4
(B) 4,2,3,1
(C) 4,2,1,3
(D) 1,2,4,3
(E) None of the above; they are all the same.

2. You put your finger over the end of a straw in a glass of water and lift the straw out of the glass. The water does not fall out of the straw. There is a small pocket of air trapped above the water in the straw by your finger. The water column does not fall out of the straw because



(A) the air in the pocket has a lesser pressure than the room's atmospheric pressure.
(B) the pressure due to the air in the pocket plus the pressure due to the water column are less than the room's atmospheric pressure.
(C) the air in the pocket has a greater pressure than the room's atmospheric pressure.
(D) the water in the straw has a greater density than air.
(E) the pressure due to the air in the pocket plus the pressure due to the water column is equal to the room's atmospheric pressure

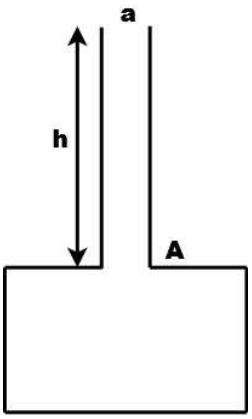
3. A narrow tube closed at one end is filled entirely with water, then inverted so that the open end is placed in a tank of water open to the atmosphere, as shown.



The water column in the tube falls to a height that is most nearly

(A) 1 m
(B) 5 m
(C) 10 m
(D) 15 m
(E) 20 m

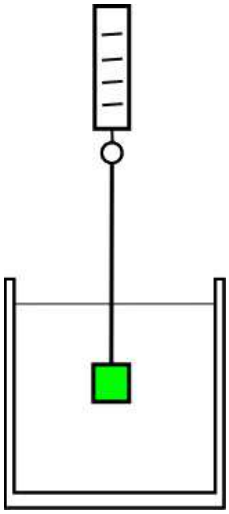
4. A pipe with area a and height h is inserted into a tank with area A, as shown.



The tank is filled with water and the pipe, which is open at the top, is filled to height h. If the radius of the tank is 0.5 m and h is 5 m, and assuming that the radius of the tank is much larger than the radius of the pipe, then the force acting on the top of the tank is most nearly:

(A) 5×10^4 N
(B) 8×10^4 N
(C) 1.2×10^5 N
(D) 2×10^5 N
(E) 3×10^5 N

5. An object of mass $m = 0.15$ kg hangs by a thread from a spring scale in a tank of water, as shown below.



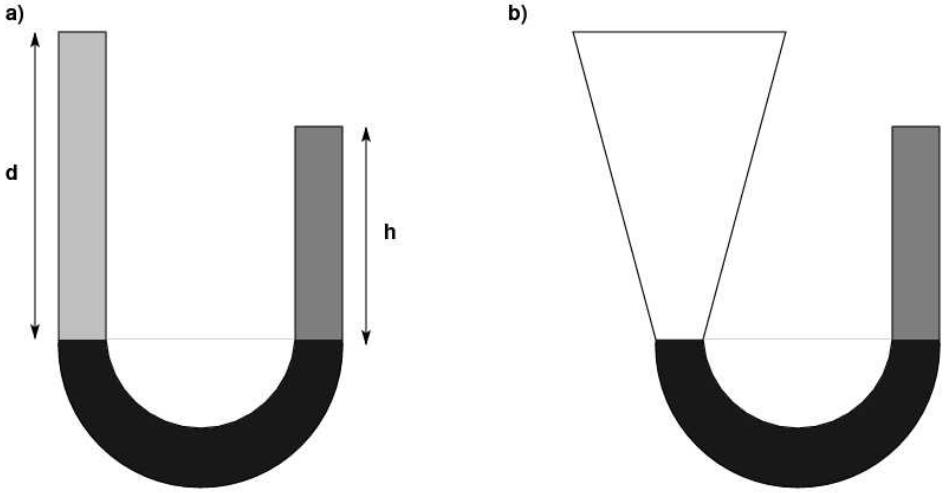
If the scale reads 0.94 N, the buoyant force is

(A) 0.15 N
(B) 0.53 N
(C) 1.5 N
(D) 2.2 N
(E) 2.5 N

6. What is the specific gravity of the object in the preceding problem?

(A) 0.53
(B) 1.89
(C) 2.77
(D) 4
(E) 5

Questions 7 - 9 refer to the following figure.



7. The U-tube on the left is first filled with mercury to a given height (bottom of Figure a, dark). It is then filled on the right with water (right-hand tube, Figure a, lighter) to a height h above the mercury, which remains at the same level. Finally, the left-hand side is filled with oil (left, Figure a, lightest) to a level d. If $h = 12$ cm and $d = 14$ cm, the density of the oil is most nearly

(A) 2.4 gm/cm^3
(B) 1.8 gm/cm^3
(C) 1 gm/cm^3
(D) 0.86 gm/cm^3
(E) 0.5 gm/cm^3

8. Assume that the left-hand tube is replaced by a funnel-shaped tube (Figure b, above). The lower U-tube is again filled with mercury to the original level. The same volumes of water and oil that were in the original configuration are now poured in to the right- and left-hand tubes respectively. What happens to the mercury level?

   (A) The mercury goes down on the left and up on the right.
   (B) The mercury goes up on the left and down on the right.
   (C) The mercury expands and goes up on both sides.
   (D) The mercury is compressed and goes down on both sides.
   (E) The mercury level stays the same.

9. You now put a piece of wood that floats on the oil in the funnel. What happens to the mercury level?

   (A) The oil-mercury column now goes down on the left and up on the right.
   (B) The oil-mercury column goes up on the left and down on the right.
   (C) The oil-mercury column expands and goes up on both sides.
   (D) The oil-mercury column is compressed and goes down on both sides.
   (E) The oil-mercury level stays the same.

10. An apparently iron cube of side length s is barely floating in water, such that it is almost entirely submerged. You suspect that there must be an air pocket inside the cube. Assuming that this is true and that the density of iron is $8\ \mathrm{gm/cm^3}$, the volume fraction of air in the cube is most nearly

    (A) 0.050
    (B) 0.275
    (C) 0.550
    (D) 0.875
    (E) 0.999

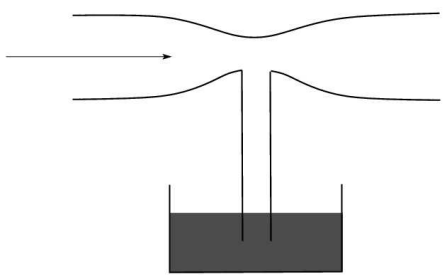
11. A coffee filter is inserted into a funnel with ridges around the inside. You attempt to remove the filter by blowing through the funnel stem. This approach fails to dislodge the paper mostly likely because

    (A) you aren't blowing hard enough.
    (B) in blowing the air you increase the pressure, which by Newton's third law causes the atmosphere to increase the air pressure against the funnel.
    (C) the blown air lowers the pressure between the filter and funnel walls, allowing the atmospheric pressure to press the filter against the funnel wall.
    (D) molecular forces cause the filter to stick to the funnel.
    (E) (C) and (D)

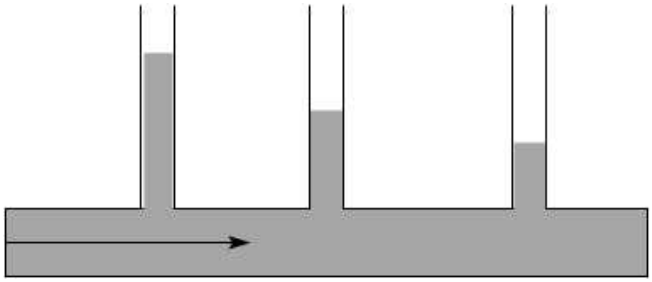
12. Air is forced through the constricted pipe, as shown in the figure below. The pan below the tube is filled with liquid open to the atmosphere.

    As the air flows through the tube,

    (A) the liquid in the pan will rise in the vertical tube.
    (B) the liquid will fall in the pan.
    (C) the level of the liquid in the pan will remain the same.
    (D) the air will force the liquid out of the pan.
    (E) whether the liquid rises or falls depends on the flow speed.



13. An incompressible liquid is flowing from left to right through a channel of constant area, as shown below. The height of the liquid in the vertical pipes drops with distance along the channel.



A possible explanation for this behavior is:

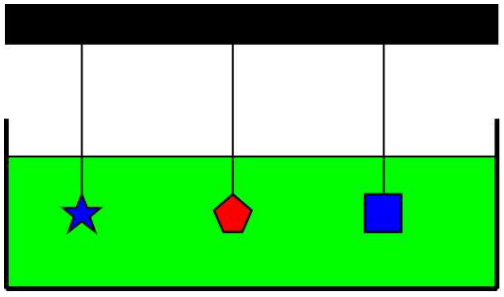
(A) the liquid is flowing faster as it moves to the right.
(B) the liquid is flowing more slowly as it moves to the right.
(C) the system is on an incline.
(D) the density is higher to the right.
(E) the pressure is higher to the right.

14. There are many explanations for the lift of an aircraft wing, some of them incorrect. However, lift is a force. Which of the following could be an expression for the lift of an aircraft wing? ($A = \text{area}; \rho = \text{mass density}; \ p = \text{pressure}; \ v = \text{velocity}; L = \text{wing length}$.)

(A) $\rho v^2 p$
(B) $\rho v^2 / L$
(C) $1/2 \ \rho v^2 L$
(D) $1/2 \ \rho v^2 A$
(E) $\rho v^2 / A$

Free-response problems

15. Three objects of identical mass but different shapes are suspended by strings in a tank of water, as shown below. All the objects are fully submerged. In this problem you may take $g = 10 \text{ m/s}^2$.



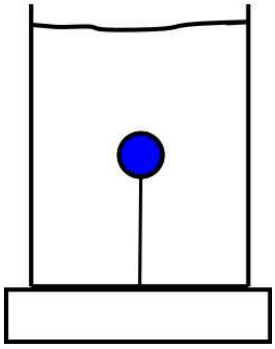
a) Must the tensions in the strings be the same? Justify your answer.

b) Assume that the cube on the right has a volume of 10^{-3} m^3 and that the tension on the string is 47.5 N. What is the buoyant force on the cube?

c) What is the cube's density?

d) If the water level in the tank is lowered such that the cube is only half submerged, what is the tension in the string?

16. A ball of mass 60 gm is tethered by a thread to the bottom of a beaker of mass 200 gm filled with 1 liter (1000 cm^3) of water, as shown. The beaker rests on a scale. The specific gravity of the ball is 0.9.



a) Draw and label the forces on the ball.

b) What is the total weight registered by the scale?

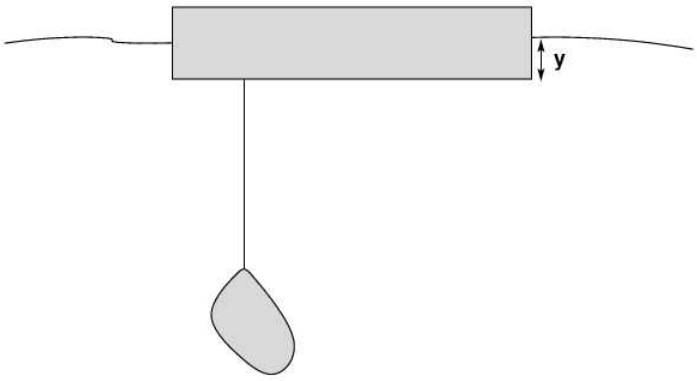
c) If the thread is cut, what is the acceleration of the ball? (Ignore viscosity and drag.)

d) What was the tension in the thread?

e) When the ball reaches the surface, what fraction is submerged?

f) Does the water level in the beaker go down, rise or stay the same? Justify your answer.

17. You have anchored your raft with a big stone tied to a rope of negligible mass, as shown below. The mass of the stone is 50 kg and its volume is 0.005 m^3. While at anchor, the raft sinks a distance y beneath the surface of the water. Take the density of water to be 1000 kg/m^3 and $g = 10\ m/s^2$.



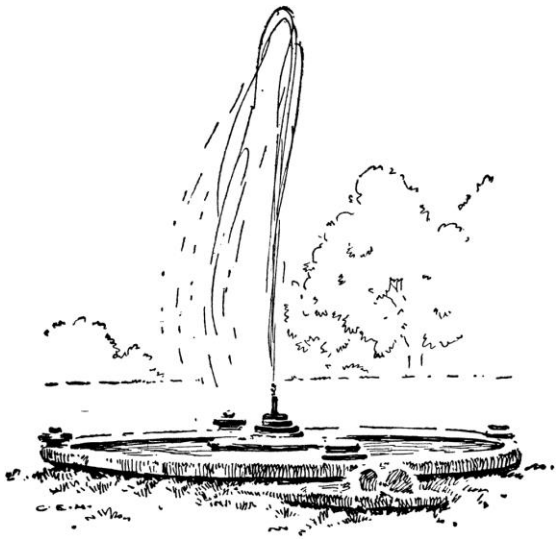
a) Draw the forces on the raft.

b) Draw the forces on the anchor.

c) What is the magnitude of the tension in the rope?

d) If the anchor is raised and placed on the raft, the raft will be submerged a distance y_1 beneath the surface of the water. Is $y_1 < y$, $y_1 = y$ or $y_1 > y$? Justify your answer. (No calculations are necessary.)

18. A fountain whose opening is 2 cm in diameter sprays water vertically into the air with a velocity of 5 m/s. You may take $g = 10\ m/s^2$.



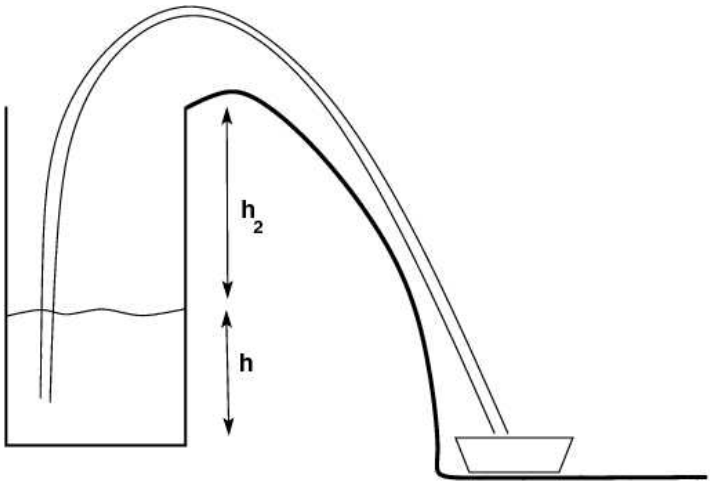
a) What is the maximum height that the water goes into the air?

b) What is the volume flow rate?

c) If the fountain is fed by a feeder pipe with diameter 3 cm, what is the absolute pressure in the feeder pipe 1.5 m below the surface of the water?

d) If the designer of the fountain wants the water to reach a height of 4 m what should he or she do without altering the volume flow rate?

19. Having tired of carrying water from a well to a container at the bottom of the hill, you run a hose from the well to the container, intending to use it as a siphon.



Let $h = 5$ m be the distance between the surface of the water in the well and the container. Let $h_2 = 8$ m be the height between the surface of the water and the top of the well. Assume the hose has a diameter of 3 cm.

a) What is the velocity of the water as it reaches the container at the bottom of the hill?

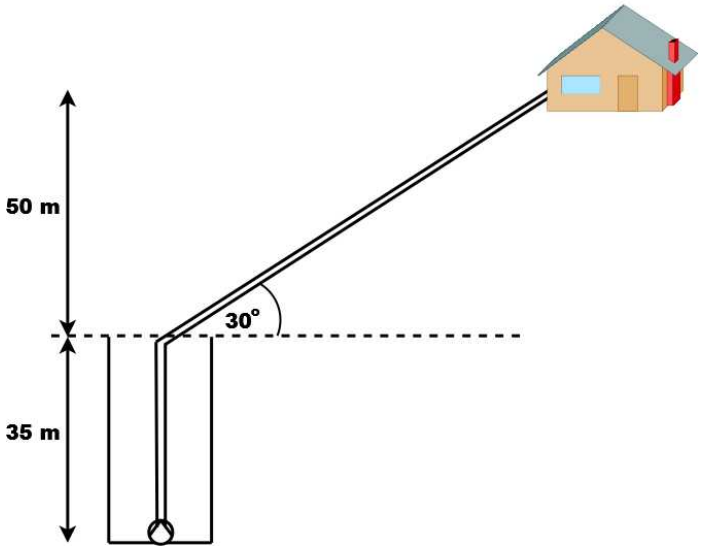
b) What is the volume flow rate?

c) What is the maximum velocity that water can drain into the container?

d) As you siphon the water out of the well, the water level drops. If the top of the well is 10 meters above the water surface, will the siphon function?

20. This is part of a problem that appeared on a 2003 AP-Physics B exam. We include it because it demonstrates how careful one must be in applying Bernoulli's principle to fluid problems.

A pump, submerged at the bottom of a well that is 35 m deep, is used to pump water uphill to a house that is 50 m above the top of the well, as shown. The density of water is 1000 kg m^3. All pressures are gauge pressures. Neglect the effects of friction, turbulence and viscosity.



The average pressure the pump actually produces is 9.2×10^5 N/m^2. Within the well the water flows at 0.5 m/s and the pipe has a diameter of 3.0 cm. At the house the pipe has a diameter of 1.25 cm.

a) Calculate the flow velocity when a faucet in the house is open.

b) Explain how you would calculate the minimum pressure at the faucet.

8. HEAT AND THERMODYNAMICS REVIEW

About ten percent of the questions on the SAT subject test concerns topics on thermodynamics. These are covered in Sections 8.1-8.5. The remaining sections are at the AP level.

1. Essential Terms

The observation that a hot object placed in contact with a cold object cools down while the cold object heats up is a basic and ancient one. To explain the phenomenon, however, turns out to be extremely difficult because although we know that something must be occurring, nothing visible is taking place. Intuitively, we say that **heat** is transferred from the hot object to the cold object, but what precisely is **heat** and how does it behave? Late eighteenth century natural philosophers in fact believed that heat consisted of an invisible fluid that moved from the hot body to the cold one. Only in the nineteenth century after Julius von Meyer and James Prescott Joule established that the potential energy released by falling weights could heat water did scientists realize that **heat is a form of energy**.

Today we recognize that **heat is exactly the energy transferred between a hot body and a cold body; it is energy due to the random motion of atoms or molecules in a substance**. However, in any solid, liquid or gas, there are such enormous numbers of particles that to calculate their motion by Newtonian mechanics is impossible. Therefore, as with the fluids of Chapter 7, physicists describe heat "in bulk," making frequent use of **pressure** (Section 7.3) and volume. To describe heat, however, an essential new ingredient must be introduced: **temperature**, which measures the kinetic energy of the molecules in a substance. The pressure, temperature and volume of a system are related to one another through what is known as an **equation of state**; the **ideal gas law** is the most famous equation of state.

Thermodynamics literally means the "dynamics of heat" and is the branch of physics that deals with heat and its behavior. Thermodynamics governs the functioning of **heat engines** and virtually the entire universe through the **first and second laws of thermodynamics**. **The first law of thermodynamics is an expression of conservation of energy that includes heat. The second law of thermodynamics is the famous law of the increase in entropy**. The basic concepts of thermodynamics, such as **temperature, specific heat, latent heat of vaporization** and **latent heat of fusion**, which appear on the SAT subject tests, are covered in Sections 1-5. More advanced concepts such as heat engines and the laws of thermodynamics, which appear on the AP exams, follow.

2. Temperature

Temperature, generally denoted T, is the measure of the average kinetic energy of an atom or molecule in a given sample of a substance. However, *temperature is not energy or any other quantity previously encountered in this book*. **Temperature is to be regarded as a new fundamental quantity along with mass, length and time**. Just as a length is what a ruler measures, practically you can think of temperature as what a thermometer measures. **Since temperature is a fundamental quantity, it has its own dimensions: temperature.** (Because many sources denote the dimensions of time by an upper case T, the dimension of temperature is frequently designated by an upper case theta, Θ.) The units of temperature are **degrees**, [∘]. Just as there are many ways to measure length, there are several temperature scales. The Fahrenheit scale, popular in the United States, is *never* employed in physics. In physics, temperature is measured either in the **Celsius** (Centigrade) **scale** or, more often, in the **absolute**, or **Kelvin scale.** The units in the Celsius scale are degrees Celsius (C) and the units in the absolute scale are **degrees absolute**, or **Kelvins** (K). **Celsius degrees are the same as Kelvins.** The only difference between the two scales is that the zero point of the Celsius scale is by definition the freezing point of water and the zero point of the Kelvin scale is **absolute zero**. As its name implies, absolute zero is meant to be the lowest possible temperature and is defined by international agreement to be $0\,\mathrm{K} = -273.15\,\mathrm{C}$. **For all problems one can take $0\,\mathrm{K} = -273^\circ\,\mathrm{C}$**. Because the degree size is the same, converting between temperature in Celsius and Kelvins is easy:

$$C = K - 273^o$$
$$K = C + 273^o \quad (8.1)$$

Thus, the freezing point of water is 0° C, or 273 K, and the boiling point of water is 100° C, or 373° K.

Perhaps the most important thing to remember about temperature is that

A system in equilibrium has only one temperature.

A hot body transferring heat to a cold body cools down while the cold body heats up. The first stops cooling and the second stops warming when they reach the *same* temperature. Thereafter, no heating or cooling takes place; any transfer of energy goes equally in both directions. We say that the system is in **equilibrium**. Furthermore, if a system A is in equilibrium with a system B and B is in equilibrium with system C, then system A is in equilibrium with system C. By the same token, *if a system is not in equilibrium, it has no well-defined temperature*. *A well-defined temperature is the defining property of thermodynamic equilibrium*.

Example 8.1: Temperature conversion. You should avoid converting from Fahrenheit to Celsius, but the conversion formula is $F = 9/5\,C\, + 32$, or $C = 5/9(F - 32)$. Thus, as you probably know, 100° C = 212° F and 0° C = 32° F. Rather than converting, you should get a feel for Celsius temperatures. The way to do this is to note that one degree Celsius is nearly twice as large as one degree Fahrenheit, $9/5$ larger to be exact. Then, remembering that 0° C = 32° F, you can immediately say that 10° C must be close to $32 + 20 = 52°$ F; the conversion formula gives 50° F. With a little practice you can even fudge the difference. It is also helpful to remember that room temperature is about 20° C and that, if you equate C and F in the conversion formula, you find that F = C at −40°. In either Fahrenheit or Celsius, −40° is cold.

Finally, many quantities that we will discuss in this chapter (for example specific heats) have values that change with temperature. Often one discusses their values at **standard temperature and pressure (STP)**, which is defined to be 1 atmosphere of pressure at 0° C. However, many texts have taken to measuring such quantities at room temperature. Pay attention to the difference.

3. Heat, Specific Heat and Latent Heat

Do not confuse heat with temperature. Temperature is a measure of the kinetic energy of a single particle. **Heat** is the *total* energy transferred from a hot body to a cold one. As a matter of daily experience, you know that it takes longer to bring to a boil a large pot of water than a small pot. This is because it takes a larger quantity of energy to raise the temperature of a large mass than a small mass. Experimentally we observe that the amount of energy necessary to raise the temperature of a substance is directly proportional to the mass of the substance and to the change in temperature:

$$Q = cm\Delta T \quad (8.2)$$

In this equation, ΔT is the change in temperature of the system and m is its mass. Q is the heat put into the system. **Because heat is energy, Q is measured in joules (J).** The constant of proportionality c is termed the **specific heat** of the substance. *Specific heat is the energy required to raise the temperature of a unit mass of a substance by one degree;* its SI units are J/(kg K). Specific heat varies from one substance to another and depends on temperature; values must be measured in the laboratory and compiled on a table. A few common values for liquids and solids are shown in Table 8.1.

Table 8.1

Substance	Specific Heat (J/kg K)	Molar Heat Capacity (J/mol K)
Lead	129	26.7
Silver	236	25.5
Copper	387	24.6
Aluminum	900	24.3
Glass	840	---
Ice (263 K)	2220	36.9
Water	4190	75.2

As you can see from the table, the specific heats of metals are small compared to those for nonconducting solids like glass, and smaller yet compared to liquids like water. Notice also that when expressed in terms of **molar heat capacity** (below), the heat capacity for all the metals is about 25 J/(mol K). There are theoretical reasons for this, which we do not discuss, but it is a useful number for your number bank.

Water has the highest specific heat of any common substance: $c = 4190$ J/(kg K), or approximately 4200 J/(kg K). Because water is so important and so exceptional in many respects, the specific heat of water is often expressed in a special unit, the **calorie**. One calorie (cal) is defined as the amount of energy required to raise the temperature of one gram of water one degree. One kilocalorie (Cal) is defined as the amount of energy required to raise the temperature of one kilogram of water one degree. Since it takes 4190 J to heat up one kilogram of water by one degree, we have 4190 J = 1 Cal and 4.190 J = 1 cal. This is known as **the mechanical equivalent of heat**. The calorie and kilocalorie are not official SI units, but in daily life they are more frequently encountered than the joule. *The calorie used when discussing food intake is actually the kilocalorie*.

When ice is heated to its melting point, 0° C, its temperature no longer increases until the ice is fully melted. Similarly, to specific heat, the **latent heat of fusion**, usually denoted L, is defined as the amount of heat necessary to melt (or heat removed to freeze) one kilogram of ice. For water L = 333 kJ/kg. Note that kJ stands for kilojoule = 1000 J. Also note that the units of L do not contain temperature in the denominator, because we are only concerned with the amount of heat necessary to melt a unit mass of ice at a *fixed* temperature. When water changes from one state (ice) to another (liquid) we say that it has undergone a **phase change**.

In the same way, the **latent heat of vaporization** is defined as the amount of heat necessary to boil one kilogram of water at 100° C. For water $L = 2256$ kJ/kg. Once again, this **phase change** takes place at a constant temperature and so there is no temperature in the expression for L.

To confuse matters, specific heats are often expressed in joules per mole per degree Kelvin (J/(mol K). As you may know, a **mole** (mol) may be thought of as an **Avogadro's number** worth of molecules.

$$N_A = 6.02 \times 10^{23}$$

That is, one mole of a substance contains 6.02×10^{23} molecules. (Molecules, not atoms. If a molecule consists of one atom, it is termed **monatomic** and the number of molecules equals the number of atoms. If a molecule consists of two atoms it is **diatomic**, and the number of molecules is half the number of atoms.) For a gas, Avogadro's number equals the number of molecules contained in 22.4 liters at STP.

Equivalently, one may think of a mole as the number of molecules in a **gram molecular weight** of a substance. For example, oxygen has atomic weight 16, but is a diatomic molecule and so its gram molecular weight is 32 gm. One mole of oxygen therefore contains 6.02×10^{23} molecules and weighs 32 grams. As mentioned, specific heat is sometimes also expressed in terms of **molar heat capacity**, which as its name implies is the amount of energy necessary to change the temperature of one mole of a substance by one degree. Molar heat capacity is measured in J/(mol K), and is also shown in Table 8.1.

Example 8.2: Boiling water. How much heat does it take to evaporate 1.5 kg of ice at $-10°$ by melting it, bringing the water to a boil, and then letting it boil away?

To melt the ice we must first raise its temperature to zero degrees. From Table 8.1, the specific heat of ice is 2220 J/(kg K). Eq. (8.2) then requires $Q_1 = 1.5 \times 2220 \times 10 = 3.33 \times 10^4$ J to raise the block of ice to its melting point. As the latent heat of fusion for water is 333 kJ/Kg, it takes $Q_2 = 1.5 \times 333 \times 10^3 = 5 \times 10^5$ J to melt it. The specific heat of water is 4190 J/(kg K), so to raise the water to 100° requires $Q_3 = 1.5 \times 4190 \times 100 = 6.28 \times 10^5$ J. The heat of vaporization of water is 2256 kJ/kg, and so to boil away the water requires $Q_4 = 1.5 \times 2256 \times 10^3 = 3.38 \times 10^6$ J. The total energy expended is $Q = Q_1 + Q_2 + Q_3 + Q_4 = 4.5 \times 10^6$ J, or about the same energy as running a 100 watt lightbulb for over 12 hours.

Example 8.3: Finding specific heat of a metal. You take a 200 gm piece of metal out of a freezer at 0° C and place it in a 1 kg pot of water at 20°, wrapping everything in a towel. Sometime later, you find that the water has cooled to 19.63° C. The metal is most likely what substance?

Assuming you have properly insulated the pot, no heat has escaped. Thus, by energy conservation the heat gained by the metal must equal the heat lost by the water, or from Eq. (8.2), $c_m m_m \Delta T_m = -c_{H2O} m_{H2O} \Delta T_{H2O}$. Once the system has reached equilibrium, everything must be at the same final temperature, $T_f = 19.63°$. Since the metal started at 0° and the water started at 20° we can write $c_m m_m (T_f - 0) = -c_{H2O} m_{H2O} (T_f - 20)$. Solving for c_m gives $c_m = -c_{H2O} m_{H2O} (T_f - 20)/m_m T_f$. Plugging in the numbers at hand yields $c_m = 395$ J/(kg K), which suggests from Table 8.1 that the metal is probably copper. However, given that you have insulated the pot with a towel, the results of this experiment shouldn't be trusted. Even a small error in the measurement of T_f will give wildly different results. (**Exercise:** Vary T_f slightly and see what happens.) This suggests why **calorimetry** experiments, as this sort of experiment is known, are difficult to reliably carry out.

4. Thermal Conductivity

Everyday experience tells you that certain materials conduct heat much faster than others. Heat one end of a metal bar on the stove, and the other end becomes too hot to touch in a matter of seconds. On the other hand, we put hot pans on stone trivets precisely because they do not allow the heat to pass. The reason metals are good conductors of heat is closely related to the reason that they are good conductors of electricity—electrons are free to move through metals, but not through insulators, like glass or stone. The free-moving electrons transfer heat from one place to another.

Intuitively, we would expect heat to travel more slowly though a long rod than a short one, and more quickly through a wide rod than a narrow one. We would also expect that the larger temperature difference from one end of the rod to the other would result in a faster transfer of heat. All these expectations are met experimentally, where one finds that the amount of heat energy transferred through a material per second (power) is given by

$$H = \frac{\Delta Q}{\Delta t} = \frac{\kappa A \Delta T}{\Delta x} \qquad (8.3)$$

Eq. (8.3) says that the power in heat transfer is directly proportional to the area A through which the heat passes, is directly proportional to the temperature difference ΔT and inversely proportional to the length of the passage Δx. (The quantity $\Delta T/\Delta x$ itself is termed the **temperature gradient**.) The constant of proportionality is κ ("kappa"), which depends on the material and is termed the **thermal conductivity**. We do not compile a table of thermal conductivities, but to give you a feel for them, the thermal conductivity for copper is $\kappa = 401$ W/(m K); for dry air $\kappa = 0.026$ W/(m K); for window glass $\kappa = 1$ W/ (m K); and for fiberglass $\kappa = 0.048$ W/(m K). One sees why fiberglass is used to insulate houses and copper isn't.

Example 8.4: Reverse engineering a soldering iron. The melting temperature of solder is typically about 200° C and a typical soldering iron is advertised at 60 W. Do the numbers make sense?

From Eq. (8.3), we require $P = 60 = \kappa A \Delta T / L$, where we set $\Delta x = L$, the length of the soldering iron tip. Let's assume that the tip is made of copper with $\kappa \approx 400$ W/(m K). The length of a soldering iron tip is about 5 cm, or 0.05 m. Soldering iron tips are fairly narrow; we'll assume a diameter of 3 mm, giving an area $A = 7 \times 10^{-6}$ m^2. Thus, $P = 400 \times 7 \times 10^{-6} \times \Delta T / (5 \times 10^{-2}) \approx 5.6 \times 10^{-2} \Delta T$. Now, $\Delta T = T_L - T_R$, where T_L is the temperature at the left end of the tip, assumed to be in the iron, and T_R is the temperature of the right end, which must be at least 200° C. With $P = 60$ W, we get $T_L = 1870°$ C.

Is that number reasonable? To heat the tip to that temperature requires an amount of heat of $Q = mc\Delta T = \rho LAc\Delta T$. The density of copper is $\rho \approx 9000$ kg/m^3. Using the same length and area as before, the specific heat of copper from Table 8.1, and assuming $\Delta T = 1850°$ above room temperature, gives $Q \approx 2260$ J. At 60 W, about 38 seconds would be required to reach the desired temperature, which seems almost plausible. However, 1870° is above the melting point of copper, which is not plausible.

Exercise: Improve this estimate.

5. Ideal Gases

Viewed macroscopically a gas is characterized by its pressure, the volume of the container holding it, and its temperature. How are these quantities related? Intuitively, it makes sense that if a piston compresses a gas, then as the volume of the gas decreases, the pressure must go up and probably the temperature as well. Indeed, for a wide variety of gases and conditions the pressure, volume and temperature are related by the famous

Ideal Gas Law

$$pV = nRT$$

$$pV = NkT \quad (8.4)$$

Any equation relating the pressure p, the volume V and the temperature T of a gas is known as an **equation of state**.[1] The ideal gas law is the most famous equation of state. We have written it in two slightly different ways. In the top version n is the number of moles present and R is the **universal gas constant**; $R = 8.314$ J/(mol K). Use the top version when you are told the number of moles of gas present. In the bottom version N is the actual number of molecules present and k is **Boltzmann's constant**; $k = 1.38 \times 10^{-23}$ J/K. Use the bottom version when you are told the absolute number of molecules in the container, for example Avogadro's number. (Indeed R is just k multiplied by Avogadro's number.)

Warning: The temperature in the ideal gas law must be expressed in degrees absolute, or Kelvins. If it is not, you will get the wrong answer.

[1] Note that if we divide both sides of the second form of the gas law by V, we get $p = (N/V)kT$. The quantity N/V is the number of particles per unit volume, or number density, call it n. The mass density ρ, mass per unit volume, is $\rho = mn$, where m is the mass of an individual particle. Thus, the perfect gas law can be written $p = (\rho/m)kT$. Any equation connecting p and ρ is also known as an equation of state. The equation $\rho = constant$, used in Chapter 7 for incompressible fluids, is the simplest equation of state.

Any gas that obeys Eq. (8.4) is termed an **ideal gas**, or a **perfect gas**. Its message is simple: **for a fixed number of molecules, the product of the volume and pressure of an ideal gas is proportional to its temperature**. On a microscopic level, any gas is a fluid in which a large number of molecules are moving around at high speeds and colliding with each other, and with the walls of the containers. An ideal gas is one in which the molecules are assumed to be: a) point particles so that they occupy zero volume, and b) interact with each other only by collisions. Except during collisions ideal gases exert no forces on each other, and in particular, no electromagnetic forces.

Example 8.5: Cooling a gas. Two moles of an ideal gas are contained in a cylinder with volume $V = 100$ liters at a temperature of 285 K. a) What is the pressure of the gas? b) If the piston in Figure 8.1 is extracted so that the volume becomes $V = 200$ liters, while at the same time the temperature decreases by a factor of two, what is the new pressure?

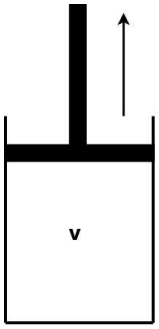


Figure 8.1

For part a), we use the ideal gas law $p = nRT/V$ for $n = 2$, $T = 285\,K$ and $V = 0.1\ \mathrm{m}^3$. Then we have $p = 4.74 \times 10^4$ Pa, which is slightly less than $1/2$ atmosphere. For part b) *do not* plug in n, R, T, the new V and recalculate the pressure. If the volume goes from V_1 to V_2, etc., the gas law tells us that $p_2/p_1 = T_2/T_1 \times V_1/V_2$. We are given that $T_2/T_1 = 1/2$ and that $V_1/V_2 = 1/2$. This immediately tells us that $p_2/p_1 = 1/4$, or that $p_2 = 1.19 \times 10^4$. **Always use proportions to solve problems when the ratio of the final quantities to the initial quantities are given. Raise your right hand and swear to this.**

Warning: Any time you use ratios in thermo problems involving temperature, use Kelvins. Although Celsius degrees and Kelvins are the same size, the ratio of two Celsius temperatures is *not* the same as the ratio of the same temperatures in Kelvins: $15/25 \neq (15 + 273)/(25 + 273)$.

6. Heat Engines and the First Law of Thermodynamics.

It is extremely helpful to recall that the subject of thermodynamics arose when nineteenth century scientists attempted to formulate the laws governing the behavior of steam engines. **The function of any engine is to perform mechanical work** (see Section 3.9). To accomplish this, the engine requires a fuel source, in the nineteenth century typically coal. As sketched in Figure 8.2, coal is burned externally and the heat released boils water in a boiler, converting it to steam. The hot steam is transferred through a pipe and an open port into a cylinder, where it drives a piston, which in turn pushes on a crankshaft attached to a train wheel. After the steam has driven the piston as far as it can, it is collected and cooled in a condenser. At the same time, a mechanism must be provided to reset the piston to its original position, in order for it to be able to drive the crankshaft a second time. This can be accomplished by a slide valve, which is itself driven such that it allows the hot steam to enter the cylinder through one port during the expansion stroke and through a port on the opposite side of the piston during the compression stroke. Meanwhile, the liquefied water in the condenser is pumped back into the boiler, where it is converted to steam again.

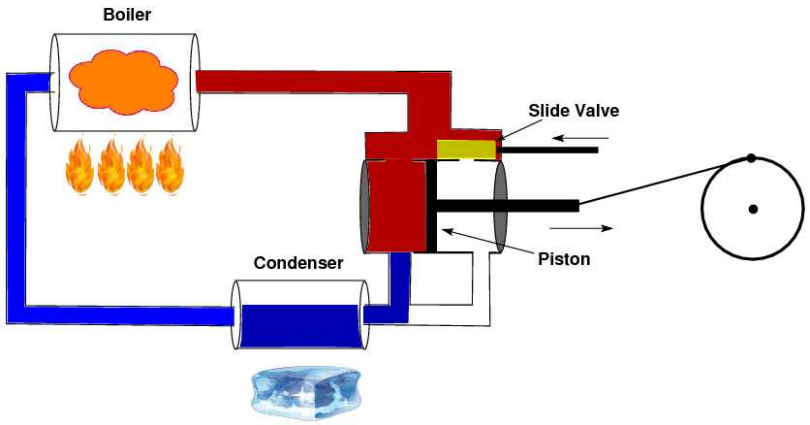


Figure 8.2

From the standpoint of basic physics there are four important features to this process: The engine requires a source of heat. Some of the heat is converted to mechanical work. Waste heat is extracted in a condenser. The process is **cyclical**, meaning that the engine is reset to its initial state so that it can repeat the operation. Any engine that operates in this general manner is referred to in physics as a **heat engine**. The source of heat is termed the **high temperature reservoir** and the coolant removing heat from the water in the condenser is termed the **low temperature reservoir**.

By conservation of energy, the heat delivered to the boiler from the high temperature reservoir, Q_H, minus any waste heat delivered to the condenser, Q_L, should be the mechanical work performed by the engine: $W = Q_H - Q_L$. Almost always one is interested in the **efficiency** of the device, which is *defined as the ratio of the work you get out to the heat you put in*, in other words, what you get versus what you pay:

$$\epsilon \equiv \frac{W}{Q_H} = \frac{Q_H - Q_L}{Q_H} = 1 - \frac{Q_L}{Q_H} \quad (8.5)$$

Notice that the efficiency can never be greater than 1 for nonzero, positive Q's. If the efficiency were greater than 1, you'd be creating energy.

The idea that you can't get something for nothing is formalized in the famous **first law of thermodynamics**. **The first law of thermodynamics is a statement of conservation of energy that includes the transfer of heat.** It is usually written

First law of thermodynamics

$$\Delta E = Q - W \quad (8.6)$$

Here, ΔE represents the change of energy of the gas itself (referred to as the **internal energy**), Q is the heat added to the gas and W is the work performed by the gas.[2] (If heat is removed from the engine, Q is negative, and if work is performed on the gas, W is negative, such that E increases.) The first law of thermodynamics is often paraphrased as "There's no free lunch," or more drastically, "You can't win."

[2] A more accurate way of writing the first law would be $\Delta E = \Delta Q \pm \Delta W$, since Q and W are increments of heat and work added to the gas. The $\pm$ serves as a warning that different texts use different sign conventions for whether the work is counted as positive or negative.

Example 8.6: Typical SAT subject test problem. A heat engine operates between two reservoirs. The high temperature reservoir absorbs 9000 J of heat at a temperature of 800 K, and the engine does 1500 J of work per cycle. How much heat is dumped per cycle into the low temperature reservoir at 300 K? What is the efficiency of this engine? In this problem, the temperatures are not relevant. If the high temperature reservoir is supplying 9000 J and the engine performs 1500 J of work per cycle, then by conservation of energy, 7500 J of energy must be lost as heat, period. From Eq. (8.5) the efficiency of this engine is $\epsilon = 1500/9000 = 1/6$, or 16.67%.

7. Further Properties of Ideal Gases

Nearly a century before the acceptance of the concept of atoms, Daniel Bernoulli proposed that gases were composed of point particles engaged in collisions with each other and the walls of any vessel containing them. Imagine the system illustrated by Figure 8.1. If a gas consists of particles moving randomly, it is reasonable to suppose that at a given instant, one-sixth of them are moving in any direction (up or down, left or right, into or out of the page). Let us then assume that at any moment one-sixth are moving upward to bang against the piston. Using the definition of force as change of momentum per unit time, we calculated in Section 3.14 that the force acting on a surface of area A due to elastic collisions with the particles hitting it is

$$F = \frac{\Delta p}{\Delta t} = \frac{\Delta p}{collision} \times \frac{collisions}{second} = 2mnAv^2 .$$

Here m is the mass of a particle, n is the number of particles per unit volume, or number density (not number of moles), and v is their velocity. The derivation assumed that all the particles were moving in the same direction. If only one-sixth are moving up, we should have

$$F = \frac{2}{6}mnAv^2 = \frac{2}{3}nA\left(\frac{1}{2}mv^2\right) .$$

By definition, the force is given by $F = pA$, for pressure p. Furthermore, the number density is $n \equiv N/V$, where N is the total number of particles present in the cylinder and V is the volume. Inserting these expressions into the above equation immediately gives

$$pV = \frac{2}{3}N\left(\frac{1}{2}mv^2\right) \quad (8.7)$$

If the speed of the particles is constant, this tells us that $pV = constant$, a fact known as **Boyle's law**, which Bernoulli's derivation explained and which was a forerunner of the ideal gas law. The quantity in parentheses is of course the kinetic energy of a particle. Then, equating Eq. (8.7) and the bottom form of Eq. (8.4) yields at once

$$KE = \frac{3}{2}kT \quad (8.8)$$

In words, *the kinetic energy of a gas particle is directly proportional to its temperature*. For this reason, we said in Section 2 that temperature is a measure of the kinetic energy of a particle in a substance. Since Boltzmann's constant is $k = R/N_A$, the total kinetic energy of N particles can be written $KE = 3/2\ (N/N_A)RT$, or

$$KE = \frac{3}{2}nRT \quad (8.9)$$

for n moles of particles.

These results are extremely significant. First, we see that Boyle's law holds when the temperature T is constant, because if T is constant, v is constant. More importantly, we see that the kinetic energy of a gas consisting of point particles interacting solely by collisions depends only on its temperature, *not* on the pressure or volume. Such a gas is termed a **monatomic ideal gas**. Because no potential energy appears anywhere—the only interactions are collisions—the only energy such a gas possesses *is* kinetic energy. For a **diatomic ideal gas**, consisting of two atoms that may rotate around an axis like a dumbbell, one must also include rotational kinetic energy (Section 4.3). Even here, however, one can show that the rotational energy also depends only on temperature. We therefore arrive at the important conclusion

The internal energy of an ideal gas depends only on its temperature

We can then rewrite Eq. (8.9) in terms of the internal energy as

$$E = \frac{3}{2}nRT \quad \text{monotonic ideal gas}$$
$$E = \frac{5}{2}nRT \quad \text{diatomic ideal gas} \qquad (8.10)$$

where we have included the expression for a diatomic ideal gas as well.

Closely related to these expressions are those for the **molar heat capacity** of ideal gases. From Section 3.9 we know that for a constant force displacing an object in the direction of the force, work = force × displacement, or $W = Fd$. If there is no displacement, *no* work is being performed. Consider the same piston in Figure 8.1 and imagine that the piston is clamped in place as the cylinder is heated. Because the piston is not moving, the gas cannot expand and can perform no work. Thus, the first law of thermodynamics, Eq. (8.6) requires that

$$\Delta E = Q.$$

The change of internal energy of the gas is merely the heat being inputted. In analogy to Eq. (8.2) we can then write

$$\Delta E = Q = c_V n \Delta T \qquad (8.11)$$

where c_v is the **molar heat capacity at constant volume**, and has the same units as the molar heat capacity for solids, $\mathrm{J/(mol\ K)}$. However, comparing Eq. (8.11) with Eq. (8.10) we see that

$$\Delta E = \frac{3}{2}nR\Delta T$$
$$c_v = \frac{3}{2}R$$
$$monotonic\ ideal\ gas$$

$$\Delta E = \frac{5}{2}nR\Delta T$$
$$c_v = \frac{5}{2}R$$
$$diatomic\ ideal\ gas \qquad (8.12)$$

If the piston *does* move, then the gas *is* performing work and the situation becomes more complicated. Let us however first assume that the pressure $p = constant$. Then for any constant-pressure process, we write in analogy to Eqs. (8.2) and (8.11)

$$Q = c_p n\Delta T \quad (8.13)$$

where c_p is the **molar heat capacity at constant pressure**. The first law, though, states that if the gas is doing work, then $\Delta E = Q - W$, and so in this case the total change of energy ΔE is due to both Q and W. By definition, $F = pA$, and so we can write the work as $W = Fd = pAd$. But Ad is just the area of the piston multiplied by the distance it has moved; it is just the change in volume of the gas ΔV, as indicated in Figure 8.3. Hence,

$$W = p\Delta V \quad (8.14)$$

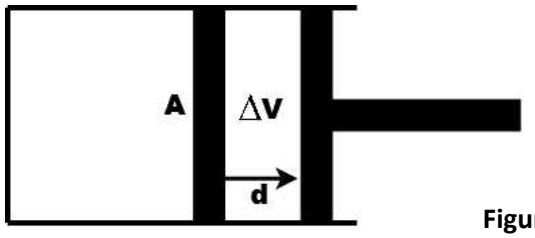


Figure 8.3

We already know that the internal energy of an ideal gas depends only on the temperature. Therefore, any change in energy must be due to a change in temperature alone. From Eq. (8.11) we can write for an ideal gas $\Delta E = c_v n\Delta T$, and thus the first law gives

$$c_v n\Delta T = nc_p\Delta T - p\Delta V$$

where we have used Eq. (8.13) for Q. But the ideal gas law itself tells us that $p\Delta V = nR\Delta T$, since everything besides V and T is constant. Substituting this into the above, and cancelling everything (**Exercise:** Do this) gives

$$c_p = c_v + R \quad (8.15)$$

Finally, one frequently has the need of a constant γ ("gamma"), which is called the **ratio of specific heats.** As its name implies, $\gamma \equiv c_p/c_v$. Since $c_v = 3/2\,R$ for a monatomic gas and $c_v = 5/2\,R$ for a diatomic gas, Eq. (8.15) gives $\gamma = 5/3$ and $\gamma = 7/5$ for a monatomic and diatomic gas, respectively.

8. Processes Involving Ideal Gases in Heat Engines

As discussed in Section 6, the purpose of a heat engine is to perform mechanical work through a cyclical process. Although real engines traditionally employed steam as the working substance, in basic physics we normally assume that the working substance is an ideal gas. We speak of the **state** of the gas as being characterized by its temperature, pressure and volume at a given point in time. The gas in a heat engine can be taken slowly from one state to another—meaning different temperatures, pressures and volumes—through any number of processes, all of which have different outcomes.

If the gas in the cylinder of Figure 8.3 is compressed or allowed to expand in such a way that the temperature remains constant, the process is termed **isothermal**. Because $pV = nRT$ and T is constant in isothermal processes, if the volume changes from V_1 to V_2 we must have $p_1V_1 = p_2V_2$. Furthermore, we established above that the internal energy of an ideal gas depends only on the temperature. *If the temperature does not change, the internal energy of the gas cannot change either.* The first law then tells us that **in an isothermal process, $\Delta E = 0$; consequently $Q = W$.** For example, in order for the gas in Figure 8.3 to expand and perform work on the piston in an isothermal process, heat must be provided, in fact exactly as much heat as the work to be done.

With calculus one can also show[3] that the work performed in an isothermal process is

$$W = nRT \ln\left(\frac{V_2}{V_1}\right) \quad (8.16)$$

where $\ln(V_2/V_1)$ is the natural logarithm of the ratio of final to initial volumes. Since $Q = W$ in an isothermal process, Eq. (8.16) also gives the amount of heat added or removed from the gas.

Any process that takes place at constant volume is termed **isovolumetric**. Because by definition there is no change in volume, Eq. (8.14) immediately tells us that **in an isovolumetric process, no work is done**. Therefore, $\Delta E = Q$. For example, if the piston in Figure 8.3 is clamped so that it cannot move, in order to change the energy of the gas one must heat (or cool) it. The change in energy of the gas is simply the amount of heat inputted. It is easy to imagine that in such a situation the pressure will rise along with the temperature. Indeed, according to the ideal gas law we have $pV = nRT$. Since the volume is being held constant we must have $T_2/p_2 = T_1/p_1$, so that as the temperature rises, so does the pressure.

Any process in which the pressure is held constant is called **isobaric**. These processes are slightly more complicated because neither the work nor the change in energy of the gas is zero; nevertheless, according to the ideal gas law, we must have $V_2/V_1 = T_2/T_1$.

Because there are various processes, all of which have different outcomes, it is extremely convenient to graph them on a ***P-V* diagram**. As its name implies, a P-V diagram plots the pressure of the gas versus its volume during a given process. What is *not* shown on a P-V diagram is the temperature, and one must infer it from the given situation, almost always by using the ideal gas law.

Example 8.7 Simple cyclical process. Figure 8.4 shows a P-V diagram for a typical cyclical process of the sort that appears on AP exams. Say 4 moles of a monatomic ideal gas is in a closed vessel with a piston. Initially the gas is in a state A at a pressure $p_o = 10$ kPa and volume $V_o = 0.2\ \text{m}^3$. The pressure in the gas is raised at constant volume until it reaches $3p_o$ at state B. The volume is next increased at constant pressure to state C where it has volume $5V_o$. From state C the pressure is reduced again at a constant volume until it reaches p_o at state D; then it is brought back at constant pressure to its original state A.

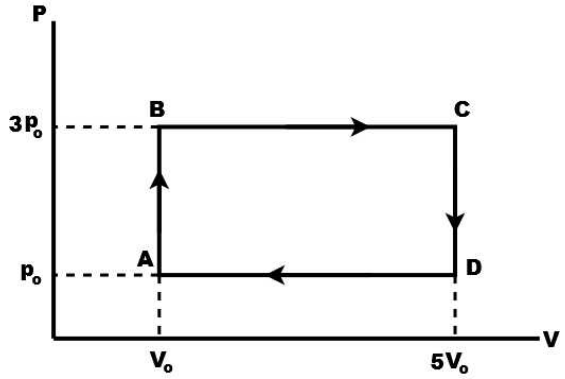


Figure 8.4

We are asked to a) rank the temperature of the four states from highest to lowest; b) decide along which paths the gas does "positive work"; c) decide whether the energy transferred to the gas along $A \to B$ is positive, negative or zero; d) calculate the net work done in one entire cycle in terms of p_o and V_o; e) calculate the efficiency of the cycle.

---

[3] By definition $W = \int_1^2 p dV$. From the ideal gas law $p = nRT/V$, with T assumed constant in an isothermal process. Therefore, $W = nRT \int_1^2 dV/V = nRT \ln(V_2/V_1)$.

It is best to work out this problem algebraically, as the problem asks, before plugging in numbers; otherwise you will become hopelessly confused. It is also best to work systematically and do the calculations one leg at a time.

a) The ideal gas law requires $T = pV/nR$. Thus, $T_A = p_oV_o/nR$. Because V is constant along $A \rightarrow B$, but $p_B = 3p_A$, we have $T_B = 3T_A$. Similarly, $T_C = 5T_B = 15T_A$ and $T_D = 5\,T_A$. Hence $T_C > T_D > T_B > T_A$.

b) and c) We are next asked to decide along which paths the gas does "positive work." This does a disservice to the student because the phrase is meaningless. There is simply no intrinsic sense to the term "positive work" or "negative work" and different texts use different sign conventions for whether work done *on* the gas or *by* the gas is counted as positive or negative.[4] The College Board should either explain their conventions or ban the term from exams. In any case, let us carry out the calculations one leg at a time:

$\boldsymbol{A \rightarrow B}$ **:** Along this path we can say immediately that the work W is 0 because $W \equiv p\Delta V$ but here the volume is constant. This corresponds to a piston being clamped, which implies that the only energy added to the system in order to raise the temperature from T_A to T_B can be heat, Q. The first law then requires $\Delta E = Q$. We know that the change of internal energy is a function of T only: $\Delta E_{AB} = 3/2\; nR(T_B - T_A) = 3/2\; nR(2T_A)$. But $T_A = p_oV_o/nR$, and so $Q_{AB} = 3p_oV_o$. This energy is added to the gas, and is generally considered "positive," in answer to (c).

$\boldsymbol{B \rightarrow C}$ **:** Along this path neither Q nor W is zero and $\Delta E_{BC} = Q_{BC} - W_{BC}$. However, the pressure is constant along this leg: $p = 3p_o$. As above we can thus write $3/2\; nR(T_C - T_B) = Q_{BC} - 3p_o\Delta V$. But $T_C - T_B = 12T_A$ and $\Delta V = 4V_o$, so $18nRT_A = Q_{BC} - 12p_oV_o$. Plugging in the above expression for T_A gives $Q_{BC} = 30p_oV_o$. This is again a positive number; heat is being added to the gas as it expands from B to C. This must be the case because the gas is doing work $W_{BC} = 12p_oV_o$ and the energy must be supplied.

$\boldsymbol{C \rightarrow D}$ **:** Once again on this vertical segment $W = 0$ because the volume is not changing. Then $\Delta E_{CD} = 3/2\; nR(T_D - T_C) = Q_{CD}$. Writing the temperatures in terms of T_A again gives $Q_{CD} = -15p_oV_o$, meaning that the gas is losing heat as the temperature decreases from C to D.

$\boldsymbol{D \rightarrow A}$ **:** As with the leg $B \rightarrow C$, nothing here is zero, and so $\Delta E_{DA} = Q_{DA} - W_{DA}$. Along this leg the pressure is again constant: $p = p_o$ and so $3/2\; nR(T_A - T_D) = Q_{DA} - p_o\Delta V$. Since $\Delta V = -4V_o$, we have $W_{DA} = -4p_oV_o$, meaning that work is being done on the gas to compress it. Once more writing the temperatures in terms of T_A gives $Q_{DA} = -10p_oV_o$.

d) Summing all the legs gives $Q_{tot} = +8p_oV_o$ and also $W_{tot} = +8p_oV_o$. That they are the same must be the case because the first law requires $\Delta E_{tot} = Q_{tot} - W_{tot}$. After a complete cycle, this "engine" is back at position A with its initial temperature T_A. Since E depends only on temperature, and after one complete cycle the temperature has gone back to its original value, so has E. Thus, $\Delta E = 0$ and Q_{tot} must equal W_{tot}.

e) The efficiency is by definition $\epsilon = W/Q$, where Q is the heat *supplied* (on legs $A \rightarrow B$ and $B \rightarrow C$). The total work is $W = 8p_oV_o$ and $Q_{supp} = 33p_oV_o$, so $\epsilon = 8/33 = 0.24$, or 24%.

Some **helpful remarks:** In doing this and similar problems it helps to keep track of the ins and outs by labeling the diagram like this:

[4] See footnote 2.

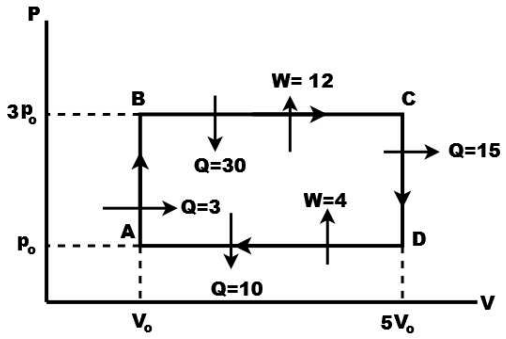


Figure 8.5

The numbers marked on Figure 8.5 are in units of $p_o V_o$. The arrows pointing into the rectangle indicate the heat added to the gas and the work done on the gas; the arrows pointing out of the rectangle indicate heat lost by the gas and work done by the gas. It doesn't matter which you call positive and negative; the only thing that matters is that the sum of the Q's equals the sum of the W's. The diagram also suggests the following shortcut.

Shortcut: Net work on a *P-V* diagram is the area within the drawn cycle. If the figure is a rectangle or a triangle, you can easily calculate the area in terms of *P* and *V*. This will be the magnitude of the net work.

However, you need to decide whether the work counts as positive or negative. In this example, note that the arrows of the cycle are going clockwise. The area of the rectangle is the total area below line BC minus the area below line DA. Below BC the area is positive because ΔV is increasing from $B \to C$. From $D \to A$ the area is counted as negative because ΔV is decreasing. In this case the sum is positive (meaning the device does net work). If you reversed the arrows, however, the result would be negative (net work is done on the gas). Thus,

Shortcut continued: The net work in a *P-V* diagram is positive (work done by gas) if the arrows point in a clockwise direction, and negative (work done on gas) if the arrows point in a counterclockwise direction.

Repeated warning: Different texts use different sign conventions for work.

The most important point to understand about this example is that in going, say, from $A \to C$, the gas ends up at the same temperature regardless of whether the path is $A \to B \to C$ or $A \to D \to C$. **The temperature, and therefore internal energy of a gas at a given position is the same, regardless of how one gets there. It is path independent.** Any variable that is path independent, is termed a **state function**. Note again, however, that the work done to get from A to C along path $A \to B \to C$ is $12 p_o V_o$, but along path $A \to D \to C$ equals $4 p_o V_o$. To make up the difference, the heat Q added to the gas along these paths is also different. Therefore W and Q are definitely *not* state functions, but via the first law they conspire to always produce the same result: ΔE. The fact that a state function is path independent implies that **the change in any state function around a closed path is zero.** We observed this to be the case for ΔE in our answer to (d).

Tip: In thermo problems, exploit the fact that ΔE is a state function when computing W or Q.

We used this trick several times in calculating Q above.

Exercise: Plug in the numbers for Example 8.7 and get the results for work and heat in joules. However, notice that nowhere did we need to do this in order to answer the questions.

Exercise: Do not assume the gas is monatomic and derive the results for Example 8.7 in terms of p_o, V_o and any constants you might need.

9. Adiabatic Processes and Free Expansion

Isothermal, isovolumetric and isobaric processes do not exhaust the inventory. If the system under consideration is totally insulated from the outside world, such that no heat is exchanged with the environment, the process is termed **adiabatic,** from the Greek *adiabatos*, "impassable." In relation to the other processes, it falls under the category "none of the above." Because no heat is exchanged with the environment, **in an adiabatic process $Q = 0$. The first law then implies $\Delta E = -W$.**

An important example of an adiabatic process is the expansion of the universe itself. Assuming that the universe is uniform from one place to another, no heat is exchanged between different regions. Thus, each part of the universe behaves as if it is insulated from its surroundings and, as it expands, its contents cool according to the adiabatic laws, below.

We state without proof that for adiabatic processes

$$pV^{\gamma} = constant$$
$$TV^{\gamma-1} = constant \quad (8.17)$$

In these expressions γ is the ratio of specific heats defined below Eq. (8.15). **Exercise:** From the first expression in Eq. (8.17) derive the second.

Note that $1/(\gamma - 1) = 3/2$ for a monatomic gas and 5/2 for a diatomic gas. Thus, for an ideal gas we can write the change of internal energy generally as $\Delta E = 1/(\gamma - 1)nR\Delta T$. With $\Delta T = T_2 - T_1$, the ideal gas law requires $p_2V_2 - p_1V_1 = nR(T_2 - T_1) = nR\Delta T$. Therefore, $1/(\gamma - 1)\,(p_2V_2 - p_1V_1) = \Delta E$. For an adiabatic process $\Delta E = -W$. Consequently,

$$W = \frac{1}{\gamma - 1}(p_1V_1 - p_2V_2) \quad (8.18)$$

Exercise: From $PV^{\gamma} = constant$ show that the above expression is positive for $V_2 > V_1$. Thus, as the gas expands it is doing work and lowering its internal energy.

Example 8.8: An adiabatic engine. Consider the thermodynamic cycle shown in Figure 8.6, below. An engine contains 1 mole of a monatomic ideal gas. The cycle begins at point A, where the pressure is 1 atmosphere and the temperature is 300 K. The gas is heated isovolumetrically until $T = 600$ K at B, then it is allowed to expand adiabatically to C, where $T = 455$ K. Finally, the gas is compressed isobarically to A again. What is the work W performed by the gas, the heat Q added and the change in internal energy ΔE along each leg? What is the efficiency of this engine?

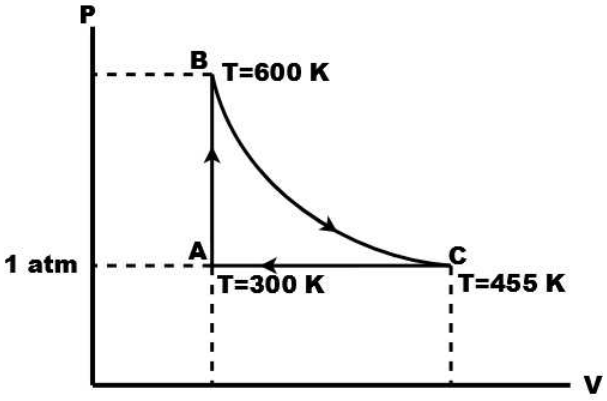


Figure 8.6

We consider each leg in turn.

$\boldsymbol{A \rightarrow B}$ **:** Note immediately that this is an isovolumetric process and so by definition along this leg the work $W = 0$. The change in internal energy can be found as before by $\Delta E = 3/2\, nR\Delta T$. With $\Delta T_{AB} = 300$ K, we get $\Delta E = 3741$ J. In this case the first law gives $Q = \Delta E = 3741$ J. This heat is being inputted into the engine in order to increase the temperature and pressure.

$\boldsymbol{B \rightarrow C}$ **:** Because this leg is adiabatic, by definition it exchanges no heat with the environment and $Q = 0$. The first law then implies $\Delta E = -W$. We again compute the change in internal energy as $\Delta E = 3/2\, nR\Delta T$. With $\Delta T_{CB} = -145$ K, we find $W = 1808$ J. This is the work done by the gas as it expands.

Exercise: Find W using Eqs. (8.17) and (8.18) and understand why you don't want to do it that way.

$\boldsymbol{C \rightarrow A}$ **:** This is an isobaric process, meaning that the pressure remains 1 atmosphere (Atm) throughout. We can thus compute the work as $W = p\Delta V$. However, we are not given the volumes. The easiest way to find them is to use the ideal gas law with $R = 0.0821$ (Atm. liter)/(mol K). Then $V_c = 37.36\ell$, $V_A = 24.63\ell$ and $\Delta V_{CA} = -12.73 \times 10^{-3}$ m^3. With 1 Atm = 1.01 kPa, we find $W = -1285$J. This is the work being done on the gas to compress it.

Again the change in internal energy is $\Delta E = 3/2\, nR\Delta T$. With $\Delta T_{CA} = -155$, we find that $\Delta E = -1933$ J. The first law says that $\Delta E = Q - W$, implying that $Q = -3218$ J, or that heat is being removed from the gas by a low temperature reservoir. (This must be the case since $T_A < T_c$.)

The net work performed by this engine is $W = 1808 - 1285 = 523$ J. We found above that the heat being added to the gas was $Q = 3741$ J. Thus, the efficiency is $\epsilon = 523/3741 = 0.14$, or 14%.

A particular circumstance that often arises in problems is termed the **free expansion** of an ideal gas, also known as **Joule expansion**. Imagine an insulated box that is partitioned into two halves with a valve between them. One half of the box contains an ideal gas and the other half contains a vacuum. The valve is opened, allowing the gas to expand into the vacuum. Because the gas is not pushing on anything, no work is being done. Furthermore, because the box is insulated, the gas is not exchanging heat with the environment and so $Q = 0$ as well. Therefore, the first law tells us that **the internal energy of an ideal gas does not change during free expansion, $\Delta E = 0$.**

Free expansion is technically what is termed an **irreversible** or **nonequilibrium** process. Roughly speaking this means that free expansion takes place quickly and the system cannot be "run backward" to its initial state in the manner of a heat engine. As mentioned in Section 2, temperature (and pressure) can be well-defined only for systems in equilibrium, i.e., that have settled down. Despite this drawback, one can use the first law to solve problems involving free expansion, as Example 8.10 in the next section illustrates.

10. The Second Law of Thermodynamics

Of all the laws of nature, the **second law of thermodynamics** is the most puzzling. At the same time it is the law that governs the real universe. There are numerous equivalent ways of stating the second law. The original form, due to Rudolph Clausius, can be phrased

Second law of thermodynamics

Heat never spontaneously flows from a cold body to a hot body

Clausius was enshrining an observation that we all know to be true: a hot object in contact with a cold object gives up heat to the cold object, thus warming it while the hot body cools. But a cold object *never* spontaneously gives up heat to a hot object, in the process growing colder itself. Yet such a process does not violate the first law of thermodynamics: a cold body *could* lose a quantity of heat to a hot body without violating energy conservation, as long as the hot body gained the same amount of energy that the cold body lost. The fact that no one had ever seen such a thing happen pointed to the need for a new law of nature.

Clausius' formulation of the second law is equivalent to an earlier observation of Sadi Carnot:

Second law of thermodynamics

It is impossible to construct a perfect heat engine

A "perfect heat engine" is one with an efficiency of 1. If it were possible to construct such a lossless device, the work it performed could be recycled back into an equivalent amount of heat, which the engine could turn into an equivalent amount of work, which could be recycled into the same amount of heat…thus creating a **perpetual motion machine**. The above statement of the second law is equivalent to saying that **a perpetual motion machine is impossible**.[5] Carnot, in fact, considered an idealized heat engine much like those of the previous section, and concluded that the maximum efficiency of such a device is

$$\epsilon = 1 - \frac{T_L}{T_H} \qquad (8.19)$$

where T_L is the temperature of the low temperature reservoir in Kelvins, and T_H is the temperature of the high temperature reservoir. Eq. (8.19) is often used for the efficiency of any heat engine and is known as the **Carnot efficiency**. (See Problem 8.23.) It is the best any heat engine can do, but notice that, for example, if $T_L = 1/2\, T_H$, the efficiency is only 50%. *The efficiency of a Carnot engine can reach 1 only when the low temperature reservoir is at a temperature of absolute zero, an impossibility*. Thus, any real-world engine must generate waste heat. If the first law implies "you can't win," a popular interpretation of the second law is, "you can't even break even."

The above formulations of the second law are in turn equivalent to the most famous version, also due to Clausius:

---

[5] Physicists actually distinguish between two kinds of perpetual motion machines. The **perpetual motion machine of the first kind** produces more energy than it takes to run, thus violating the first law of thermodynamics. The **perpetual machine of the second kind** is perfectly efficient and violates the second law of thermodynamics. Both are impossible.

Second law of thermodynamics

In any process the entropy of an isolated system never decreases

To understand this statement one must of course have a definition of **entropy**. Before that, however, we distinguish between two types of processes: **time-reversible** and **time-irreversible**. A **time-reversible**—or **reversible** process for short—is one that can retrace its steps. Reversible processes take place slowly; all changes are assumed to be very slight so that they can be "undone." The oscillations of an idealized spring constitute a reversible process, because the spring can and does retrace every motion it makes. The idealized cyclic engines in the previous section were reversible: running them backward brings them exactly through the states they passed through while running them forward, such that the temperature and pressure at each point are the same; moreover, the amount of heat expelled on each leg is the same as was previously absorbed, etc. At each step in a reversible process the system is assumed to be in equilibrium, so that it has a well-defined temperature and pressure.

For elementary physics, the most important point is that reversible processes take place *only in the absence of friction.* Friction impedes motion while generating heat and prevents a spring from exactly repeating its oscillations, or an engine from returning exactly to its original state. All processes in the real world are of this type: **time-irreversible**. **A perfectly reversible process is not achievable in the real world.**

Entropy is defined as the heat absorbed by a system divided by the temperature: $S = Q/T$. Generally, we are more interested in changes in entropy, and for small increments one can say

$$\Delta S = \frac{\Delta Q}{T} \qquad (8.20)$$

Heat, Q, is always exchanged, and is always a "delta," but we write ΔQ for emphasis. The second law states that **for any process in an isolated system, $\Delta S \geq 0$**. *The equality holds only for reversible processes*. For anything that takes place in the real-world, entropy always increases. In Eq. (8.20), T must *always* be expressed in Kelvins.

Example 8.9: Melting ice. An ice cube of 100 gm melts slowly in a pan of water at 0° C. What is the change in entropy of the system? From Section 3, the latent heat of fusion for water is 333 kJ/kg. To melt a 100 gm ice cube we therefore require 3.33×10^4 J. From Eq. (8.20) the change in entropy is thus $\Delta S = S_{H2O} - S_{ice} = 3.33 \times 10^4/273 = 122$ J/K. Note that this is a positive number merely because we assumed that $Q = +3.33 \times 10^4$ J was being added to the ice cube from the reservoir of water, the pan. The ice cube in this problem is *not* an isolated system because it is in contact with the reservoir. The "isolated system" in this case is the cube *plus* the pan, which is assumed not to interact with the rest of the universe, and the words "melts slowly" is exam code for "reversible process." The change of entropy of the entire isolated system is just zero because the reservoir is giving up exactly as much heat as the ice cube is absorbing. In real life this would not be the case and ΔS would be greater than zero.

Because the second law states only that $\Delta S > 0$ in irreversible processes, it can be difficult, if not impossible, to calculate the exact value of the change in entropy. Nevertheless, an important property of entropy allows such calculations to be carried out in relatively simple situations: **Entropy is a state function.** The value of entropy at any point in a process is unique. Like internal energy, it is path independent. In Figure 8.4, for example, the entropy at point C has one and only one value. It does not matter whether you reach C along the path $A \rightarrow B \rightarrow C$ or $A \rightarrow D \rightarrow C$. This fact allows us to calculate entropy changes in certain irreversible processes.

Example 8.10: Free expansion of an ideal gas. Suppose that 1 mole of an ideal gas expands freely from a volume V to a volume $3V$ at a temperature of 293 K. What is the change in entropy? Free expansion as discussed in in Section 9is an irreversible process; it takes place too rapidly for the system to reach equilibrium and so it has no clear temperature. Nevertheless, we know that in free expansion the temperature of the gas does not change; it is therefore an isothermal process. Eq. (8.16) gave the work performed in isothermal expansion. Although this was technically for a reversible process we can use it to calculate the entropy change in free expansion because entropy is a state function. That fact allows us to compute ΔS between two states in the most convenient manner possible; any chosen path must give the same answer. For isothermal expansion, the first law gives $\Delta E = 0 = Q - W$, implying that $Q = W$. Consequently, from Eq. (8.17), $Q = RT \ln(V_2/V_1)$ and the change in entropy is $\Delta S = R \ln(3) = 9.13$ J/K.

11. Interpretations of Entropy

Entropy is a difficult concept and this has led to many statements that are only partially correct. For example, entropy is almost universally equated with randomness, disorder, or at least labeled as a measure of a system's disorder. When a fresh egg in its shell splatters on the floor, we definitely feel that disorder has increased. Nevertheless, such an interpretation is not always correct. If one end of a sealed pipe containing two different types of molecules is heated, the two species will separate. In the process, collisions will increase entropy, but the final state, with the two species at opposite ends of the pipe, is more ordered than the initial state, in which the two types of molecules were mixed.

The second law is puzzling above all because it is the only law of nature that requires a certain quantity to always *increase*. The other laws, like conservation of energy or momentum, refer to the *equality* of a quantity before and after a process takes place. For this reason, the second law is unique among the laws of nature in providing an **arrow of time**. The egg splattering on the floor is certainly an irreversible process; we never see a splattered mess on the floor reconstitute itself into a whole egg, unless it is in a movie run for comic effect and we know that it is run backwards. Thus, time-irreversible processes, as their name implies, proceed in only one direction in time. All the other laws, in particular $F = ma$, are time reversible—one cannot tell whether a film of, say, the planets in orbit around the sun is being run forward or backward.[6] But it is the irreversible processes that govern the real world, result in the increase of entropy and distinguish past from future. To reconcile the second law with the other laws of physics, in particular to explain how time-irreversible processes can arise from time-reversible laws, has a long and controversial history. Unfortunately, space does not permit us to enter into such matters here.

THERMODYNAMICS PROBLEMS

Questions 1-10 are at the subject-test level. The remaining problems are at the AP level.

1. A block of copper with mass 0.5 kg is in thermal equilibrium with a second block of copper of mass 2 kg, which is in thermal equilibrium with a block of aluminum of mass 5 kg and temperature 300 K. The temperatures of the 0.5 kg copper block and the 2 kg copper block are, respectively

   (A) 300 K and 300 K
   (B) 150 K and 600 K
   (C) 600 K and 150 K
   (D) 75 K and 300 K
   (E) 300 K and 75 K

[6] The only place time appears in $F = ma$ is in the denominator of the acceleration: $a = d^2x/dt^2$, but here the time interval is squared, so it makes no difference whether it is positive or negative.

2. If it takes 1.30 kJ of heat to change the temperature of a 1.5 kg block by 5°, then its specific heat is most nearly:

   (A) 23 J/(kg K)
   (B) 55 J/(kg K)
   (C) 173 J/(kg K)
   (D) 5012 J/(kg K)
   (E) 2213 J/(kg K)

3. A block of aluminum at a temperature of $T = 100°$ C is placed in an insulated container of water at $T = 20°$ C. After reaching equilibrium

   (A) the temperature of the aluminum will be less than the temperature of the water.
   (B) the temperature of the aluminum will be greater than the temperature of the water.
   (C) the equilibrium temperature will be greater than 100° C.
   (D) the temperature of the aluminum and water will be the same and between 20° C and 100° C.
   (E) None of the above

4. Two identical masses of water are cooled ten degrees in two identical freezers. The first mass of water is cooled from 25° C to 15° C and the second mass of ice is cooled from 5° C to −5° C. One can say that

   (A) the amount of heat extracted from mass one is greater than the amount of heat extracted from mass two.
   (B) the amount of heat extracted from mass two is greater than the amount of heat extracted from mass one.
   (C) it should take longer to cool mass two than mass one.
   (D) you don't know how much heat is extracted unless you know the work performed by the freezers.
   (E) (B) and (C)

5. Liquids A and B are identical except that liquid A has a boiling point of 60° C, liquid B has a boiling point of 120° C and the latent heat of vaporization of A is half that of B. Let $m_A = 2.5\, m_B$, where m_A and m_B are the masses of A and B, respectively. Raising the temperature of each liquid from 20° C on identical stoves, one can predict that

   (A) the two liquids will take the same amount of time to boil away.
   (B) A will boil away sooner than B.
   (C) B will boil away sooner than A.
   (D) One can't make a firm prediction without knowing the latent heats of vaporization.
   (E) One can't make a firm prediction without knowing the specific heats.

6. A block of aluminum is cooled in a refrigerator, then placed in a jug of water. In order to calculate the specific heat of aluminum, you must know all of the below quantities *except*

   (A) the specific heat of water.
   (B) the masses of the aluminum and water.
   (C) the original temperatures of the aluminum and water and their final, equilibrium temperature.
   (D) the time necessary for the system to reach equilibrium.
   (E) the mass of the aluminum.

7. A 1 kg block of aluminum at a temperature of $T = 100^\circ$ C is placed in an insulated vessel containing 5 kg of water at $T = 20^\circ$ C. The specific heat of aluminum is 900 J/(kg K) and the specific heat of water is 4190 J/(kg K). After reaching equilibrium

   (A) the final temperature will be about 50° C.
   (B) the temperature of the aluminum will be 25.7° C and the water temperature will be 23.9° C.
   (C) the final temperature will be about 23° C.
   (D) the temperature of the aluminum will be 23.8° C and the water temperature will be 30.1° C.
   (E) the final temperature will be about 75° C.

8. Pizza in an oven tends to heat rapidly compared to a food X of equal mass, but also cools rapidly once it is taken out of the oven. From this observation you might conclude that

   (A) the specific heat of pizza is very high compared to the specific heat of X.
   (B) the specific heat of pizza is very low compared to the specific heat of X.
   (C) the density of pizza is lower than the density of food X.
   (D) the density of pizza is higher than the density of food X.
   (E) None of the above

9. You come across a stone brick and a piece of metal lying on the ice on a cold day. The metal feels colder than the brick, although they are at the same temperature. The best explanation for this phenomenon is that

   (A) the thermal conductivity of metal is greater than that of brick.
   (B) the thermal conductivity of metal is less than that of brick.
   (C) the latent heat of evaporation of the metal is greater than that of the brick.
   (D) the density of metal is less than that of brick.
   (E) the density of metal is greater than that of brick.

10. Two bars of identical dimensions have been sticking out of the same furnace for a long time. You know the specific heats of the bars, the thermal conductivity, their masses and the temperature of the ends sticking out of the furnace. To make a plausible determination of the furnace's temperature, you could use

    (A) the specific heat of the bars alone.
    (B) the specific heat of the bars and their masses.
    (C) the thermal conductivity of the bars alone.
    (D) the thermal conductivity of the bars and the temperatures of the ends protruding from the furnace.
    (E) Either (B) or (D).

11. If a fixed amount of gas in an engine does 62 J of work and loses 30 J of heat to the environment, then the internal energy of the gas

    (A) gains 92 J
    (B) loses 92 J
    (C) remains the same
    (D) loses 30 J
    (E) loses 62 J

12. If a heat engine is supplied with 560 J of heat per cycle, loses 420 J to the environment and does 140 J of work, then its efficiency is

    (A) 14.3%
    (B) 26.9%
    (C) 33.3%
    (D) 75.0%
    (E) None of the above

13. If the condenser of a steam engine is at 20° C and the high temperature reservoir is at 120° C, then the maximum amount of work you can get out of 1000 J of heat supplied to the engine is most nearly

    (A) 1000 J
    (B) 833 J
    (C) 330 J
    (D) 170 J
    (E) 60 J

14. The heat of combustion produced by gasoline is approximately 47 kJ/mol. If a gasoline engine operates between 1500 K and 750 K, then the maximum height that 5 moles of gasoline could lift a 400 kg aircraft is most nearly

    (A) 10 m
    (B) 20 m
    (C) 30 m
    (D) 40 m
    (E) 50 m

15. An amount of heat Q is delivered from the left end of a rod of length L to the right end in a time Δt. The cross-sectional area of the rod is A, the left end of the rod is held at a temperature T_L and the right end at a temperature T_R, with $T_L > T_R$. To be certain of increasing the amount of heat delivered, you could

    (A) increase L only.
    (B) increase A and increase L.
    (C) increase A, decrease L and increase $T_L - T_R$.
    (D) decrease L and decrease $T_L - T_R$.
    (E) decrease A only.

16. A fixed amount of a perfect gas in a cylinder expands from V to $2V$, in the process doing work against a piston. If the pressure is held constant, one can conclude which of the following: i) the temperature of the gas increases; ii) the internal energy of the gas increases; iii) the temperature of the gas goes down; iv) the internal energy of the gas goes down; v) heat is lost from the gas; vi) heat must be added to the gas?

    (A) i and ii
    (B) iii and iv
    (C) iii, iv and v
    (D) vi
    (E) i, ii and vi

17. Two identical sealed containers, each of volume V contain ideal gases, either monatomic or diatomic. It is observed that 90 J are required to heat the gas in container A ten degrees and 150 J are required to heat the gas in container B ten degrees You can conclude that

    (A) container A holds a monatomic gas and container B holds a diatomic gas.
    (B) container A holds a diatomic gas and container B holds a monatomic gas.
    (C) the two containers hold the same number of moles of gas.
    (D) the two containers hold a different number of moles of gas.
    (E) None of the above

18. If the average molecular speed of the molecules in a container is doubled, then the pressure that the gas exerts on the vessel walls should

    (A) stay the same.
    (B) go down by $\sqrt{2}$.
    (C) go up by $\sqrt{2}$.
    (D) double.
    (E) go up by 4.

19. If the temperature of the gas in a container is halved, then the average molecular speed should

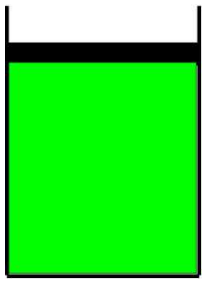
    (A) stay the same.
    (B) go down by $\sqrt{2}$.
    (C) go up by $\sqrt{2}$.
    (D) double.
    (E) go up by 4.

20. If an ideal gas is taken from a state A around a complete thermodynamic cycle to state A again, quantities that return to their initial values are: i) the temperature of the gas; ii) the work performed by the gas; iii) the work performed on the gas; iv) the heat exchanged with the environment; v) the internal energy of the gas.

    (A) i
    (B) v
    (C) i, ii, iii
    (D) i,ii, iii, iv
    (E) i, v

Free-response problems

21. An ideal gas is contained in a cylinder by a piston, as shown in the figure. The top of the cylinder is open to the atmosphere; the piston has mass m and area A and is free to move up or down. Initially, the system is in equilibrium and the volume of the cylinder is V_o.



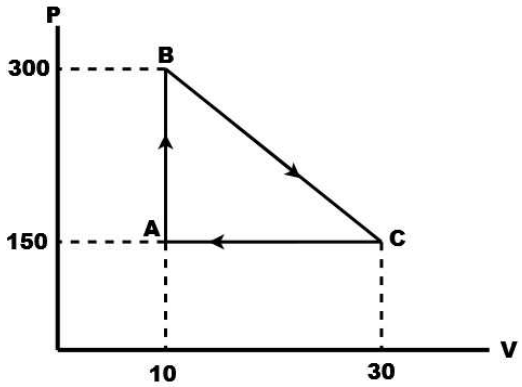
a) What is the pressure that the gas exerts on the piston?

b) The piston is clamped in position and an electric heater is placed under the cylinder. If the power of the heater is P and one assumes no heat loss, after a time Δt, how much heat has been delivered to the gas in the cylinder?

c) The piston is unclamped and the gas is allowed to expand adiabatically to a volume V_f. Does the temperature increase, decrease or stay the same during this expansion? Justify your answer.

d) The gas is then compressed isothermally to its original volume, V_o. Draw a P-V diagram for the entire cycle.

22. One mole of an ideal gas is taken through the thermodynamic cycle shown below, from $A \to B \to C \to A$. The indicated pressures are in $\mathrm{N/m^2}$ and the volumes are in $\mathrm{m^3}$.



a) What is the temperature at points A, B and C?

b) What is the work carried out during one cycle?

c) Is net heat added to the gas or removed from the gas?

d) Has the entropy of the gas changed after one cycle? Justify your answer.

e) Has the internal energy of the gas changed after one cycle? If so, how much? Justify your answer.

f) Which of the previous responses change if the engine is run backwards, from $A \to C \to B \to A$?

23. The most efficient heat engine conceivable is a perfectly reversible engine known as the **Carnot engine**, or **Carnot cycle**, after Sadi Carnot. The Carnot cycle consists of four steps:

    1) An ideal gas initially at a volume V_A expands isothermally to volume V_B.
    2) From V_B, the gas expands adiabatically to a volume V_c.
    3) From V_C the gas is compressed isothermally to volume V_D.
    4) From volume V_D the gas is compressed adiabatically until it reaches V_A, completing the cycle.

    a) Sketch a P-V diagram for the Carnot cycle. Justify how you've drawn each leg.

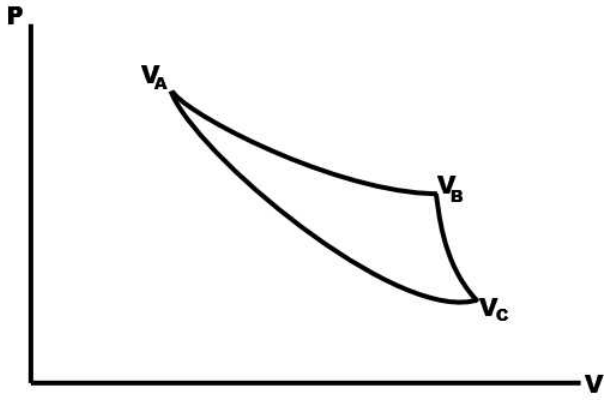
    b) Mark where heat is being absorbed by the gas and removed from the gas, and also where the work is being done by the gas or on the gas.

    c) Calculate the net work done by this cycle in terms of T and V.

    d) Calculate the efficiency of this engine in terms of temperature alone.

    e) Explain in your own words why it is the most efficient engine possible.

24. A three-step cyclical process is shown below:



    1) An ideal gas initial at a volume V_A expands isothermally to volume V_B.

    2) From V_B, the gas expands adiabatically to a volume V_c.

    3) From V_C the gas is compressed adiabatically to volume V_A again, completing the cycle.

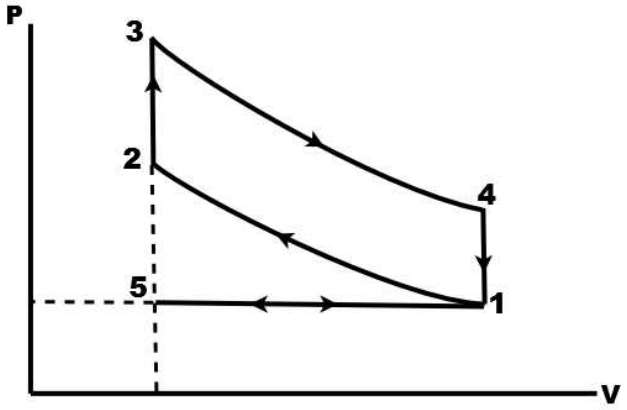
    a) Mark where heat is being absorbed by the gas and removed from the gas, and also where the work is being done by the gas or on the gas.

    b) Calculate the net work done by this cycle in terms of T and V.

    c) Calculate the efficiency of this engine in terms of temperature alone.

    d) Explain in your own words why this engine is impossible.

25. The **Otto cycle** is the thermodynamic cycle that most closely describes an actual car engine. It consists of the following steps as shown on the below *P*-*V* diagram:



1) An intake stroke, $5 \rightarrow 1$, in which the fuel mixture is drawn into the cylinder.

2) An adiabatic compression stroke, $1 \rightarrow 2$, in which the fuel mixture is compressed.

3) Ignition by the spark plug, followed by combustion. This takes place in a short time, essentially at constant volume, $2 \rightarrow 3$.

4) An adiabatic power stroke, $3 \rightarrow 4$, in which the expanding gas does work.

5) Valve exhaust: A valve opens, the gas is exhausted, heat escapes and the temperature is lowered, $4 \rightarrow 1$.

6) Exhaust stroke: the piston is compressed at constant pressure, $1 \rightarrow 5$, expelling the remaining combustion products from the cylinder.

a) Mark on the diagram the places where heat is added to or removed from the gas.

b) Calculate the net heat added to the gas in terms of the temperatures and c_V, the specific heat of the gas at constant volume, assuming an ideal gas.

c) Calculate the efficiency of the engine in terms of the temperatures alone.

d) Calculate the efficiency in terms of the volumes.

e) Calculate the efficiency in terms of the compression ratio $r \equiv V_1/V_2$.

9. ELECTRICITY AND MAGNETISM REVIEW

The material covered by this chapter is essential for both the SAT subject test and the AP exams. Sections 9.1-9.5, 9.7-9.9, 9.11-9.12 are at the subject-test level. Sections 9.6, 9.10, 9.13 and 9.14 are at the AP level.

1. Essential Terms

The previous chapters have dealt almost exclusively with "classical," or Newtonian, mechanics, which concerns the three fundamental quantities of mass, length and time. While mechanics was developed in the seventeenth century, the subjects of **electricity** and **magnetism** did not receive formal development until 1820, when Hans Oersted observed that **electric currents** deflected compass needles, immediately establishing that the two phenomena are inseparable. Consequently, the subject of electricity and magnetism is often referred to as **electromagnetism**. All the laws of nature previously discussed remain in force. The main distinction between electromagnetism and Newtonian mechanics is the introduction of a new quantity, **electric charge**, which in elementary physics is taken to be as fundamental as mass, length and time. The existence of electric charge gives rise to **electrical and magnetic forces and fields, electrical currents** and a whole range of phenomena that underlie modern civilization. In terms of the SAT, over twenty percent of the problems on the subject test concern electromagnetism and often an even higher percentage on the AP exams.

2. Electric Charge

The ancient Greeks understood that rubbing amber with fur results in an **electric charge**. English remembers this: our word "electron" comes from the Greek $\eta\lambda\varepsilon\kappa\rho\upsilon$—"electron," meaning amber. Electric charge, usually denoted q, plays a role analogous to mass m in mechanics, except that **electric charge comes in two types, termed positive and negative, or $+$ and $-$. The most basic observed property of charge is that like charges repel; unlike charges attract.** Two positive or two negative charges repel each other, while a positive and a negative charge attract each other.

Like matter, charge obeys a fundamental, invincible law of nature:

Conservation of Charge

The total amount of charge in an isolated system cannot change.

Given that the total amount of charge in a system is fixed, what does it mean to "charge" an object? To "charge an object" means merely to separate any charges present on it, such that there is an excess of positive charge in one place and negative charge in another place. However, *if there is initially zero charge on the object, the total charge remains zero*.

Mass in Newtonian physics can be divided into infinitesimally small quantities. Unlike mass, there exists a smallest amount of electric charge, which we take to be the charge on an **electron** or **proton** (the charge on the electron is usually denoted e^- and is negative, while the charge on the proton has the same magnitude but is positive; it is usually denoted e or p^+). The fact that there is a smallest unit of charge means in modern language that charge is **quantized**, or comes in discrete amounts (Section 12.6). The official SI unit of charge is, however, the coulomb (C). In terms of the charge on the electron or proton:

$$1\ \text{C} \equiv 6.24 \times 10^{18}\ e$$
or
$$e = 1.6 \times 10^{-19}\ \text{C} \quad (9.1)$$

The total charge of a system (often denoted Q) counts the number of elementary electrical charges on it; Q must therefore always be an integral multiple of e.[1]

Tip: A coulomb of charge is an extremely large number and rarely encountered in typical situations. More often charge is measured in microcoulombs ($1\ \mu\text{C} = 10^{-6}$ C) or nanocoulombs ($1\ \text{nC} = 10^{-9}$ C). A lightning strike carries roughly 10 C of charge.

Example 9.1: Pennies. Modern pennies are made of zinc because to mint a copper penny costs more than a penny. Suppose one out of every 10,000 zinc atoms in a penny has an electron removed. What is the resulting charge on the penny?

A penny has a mass of 2.5 gm and the atomic weight of zinc is approximately 65 gm. Thus, a penny represents $2.5/65$ or 0.38 moles of zinc. Multiplying by Avogadro's number 6.02×10^{23} gives 2.3×10^{22} atoms. If one out of every ten thousand of these is stripped of an electron, then the number of electrons taken away is 2.3×10^{18}, about a third of a coulomb. By conservation of charge, the number of positive **ions**, or charged atoms, remaining on the penny is the same.

3. Charging Objects

Because charge is conserved, "charging an object" means only to move charges around from one place to another. There are several ways to accomplish this. The first is **charging by contact**. As its name implies, to charge an object by contact means to touch it to an already charged object. The uncharged object will then pick up some of the charges from the charged objects. For example, suppose an uncharged, or **electrically neutral**, sphere A is placed in contact with an identical sphere B that carries charge 10 μC. Assuming that the charges are evenly distributed, when sphere A is removed from sphere B, they will each carry 5 μC of charge.

The charges will tend to redistribute evenly if the spheres are **conductors**. **A conductor is a material that allows electrons to move freely**. Typically, metals are good conductors and non-metals, such as glass, plastic, rubber and stone, are poor conductors, or **insulators**. As mentioned in Section 8.4, good electrical conductors also tend to be good conductors of heat, and for much the same reason—the electrons are free to move, transferring energy from one location to another.

A second method for charging objects is termed **charging by induction**, which means to charge an object across space without actually touching it. Suppose a positively charged rod is brought near an electrically neutral metal sphere, as in Figure 9.1a.

[1] Lower-case q tends to be used for single charges and Q for either total charges or source charges, as we will, but there is no fixed rule.

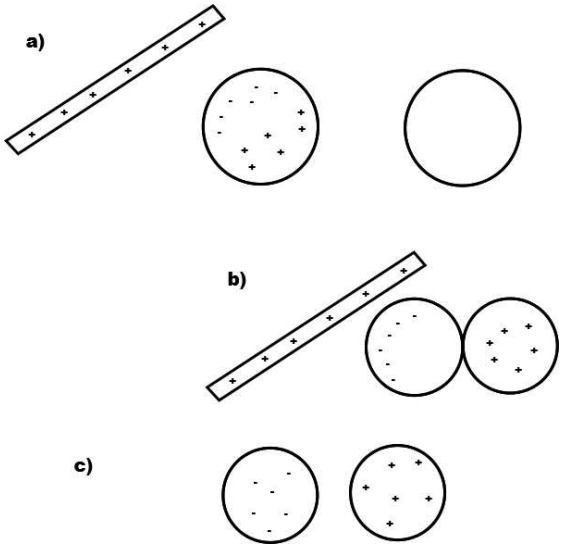


Figure 9.1

The charges on the sphere will separate such that the negative charges are attracted to the rod and the positive charges are repelled by it. We say that a charge has been **induced** on the sphere, although at this point the charges are merely separated. When charges are separated on an object we also say that the object has been electrically **polarized**. If a second, identical, neutral sphere is now touched to the first sphere, the positive charges will migrate to the second sphere, giving it a positive charge $+Q$, as in Figure 9.1b. If the two spheres are then separated, the first sphere will carry a charge $-Q$ and the second sphere will carry an equal and opposite charge $+Q$, as in Figure 9.1c.

4. Coulomb's Law

If two charges are attracting or repelling each other, this is a push or a pull. Consequently, the charges must be exerting a force on each other. The law governing the electrical force between two single charges is known as

Coulomb's Law

$$F = \frac{kq_1q_2}{r^2} \qquad (9.2)$$

after Charles-Augustin de Coulomb. Eq. (9.2) gives only the magnitude of the force between the charges but does not tell you its direction. Note immediately and for all time that Coulomb's law is identical in form to Newton's law of gravity, Eq. (3.4). Like the law of gravity, the strength of the force depends on the product of the two charges, q_1 and q_2. Like the law of gravity, its *most important* feature is that it is an **inverse-square law**; if the distance r between the two charges doubles, the force decreases by a factor of four.

Like gravity, the Coulomb, or **electrostatic**, force acts along the line joining the centers of the two charges, which are assumed to be single points or spheres. Unlike gravity, which is always attractive, **the Coulomb force between charges can be attractive or repulsive, depending on the sign of the charges.** The constant k, which determines the size of the force, must be measured in the laboratory and is found to be $k = 8.99 \times 10^9\ \text{N m}^2/\text{C}^2$. *For almost all purposes you can take* $k = 9 \times 10^9$.

Even without knowing k, one can still draw certain conclusions from the form of Coulomb's law alone.

Example 9.2: Forces on a test charge. Suppose a positive **test charge** q lies directly between two other charges q_1 and q_2, but closer to q_1, as shown in Figure 9.2. If q does not move, what can we conclude about q_1 and q_2?

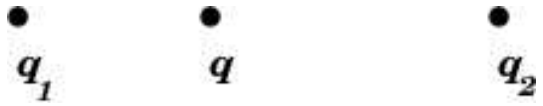


Figure 9.2

First, a **test charge** is merely the designated charge *on which* the forces under consideration are acting. It is completely analogous to the m in $F = ma$. By convention **a test charge is assumed to be positive unless otherwise specified**. Since q isn't moving, we can immediately say that the total force acting on it by q_1 and q_2 must be zero. That in turn implies that the forces q_1 and q_2 exert on q must be in opposite directions, in order to cancel out. Thus, q_1 and q_2 must either be both positive or both negative. In the first case they will be pushing on q in opposite directions; in the second case they will be pulling on q in opposite directions. (**Exercise:** Draw in arrows showing the direction of the forces q_1 and q_2 exert on q.) Both cases are allowed if q is stationary. Moreover, we can also conclude that $q_2 > q_1$. According to Coulomb's law, the electrostatic force diminishes with the square of the distance; because q_2 is farther away from q than q_1, it must be larger than q_1 to compensate for the extra distance.

Exercise: What changes if q is negative?

5. Electric Fields

The notion of a **field** was originally conceived by natural philosophers to explain a great mystery: The mechanical forces of the type discussed in Chapter 3 are all transmitted by direct contact between objects. How then could gravitational and electric forces be transmitted across empty space? The mind boggled. The **field** was introduced as the medium that transmits a force across a distance. In terms of the electrostatic force, call the source charge Q. If it exerts a force *on* a test charge q, then the electric field is transmitting the force *from* Q *to* q, and so the field should belong to Q alone. After all, Q's field might affect other test charges as well, and so it should not depend on them. Although the original motivation for the concept of field as a transmitter of force is somewhat outdated, **fields are real and can be measured**. What remains of the original idea is that the field depends on only the *source*. The electric field of a point particle is termed the **Coulomb, or electrostatic, field** and for a source charge Q its magnitude is defined as

$$E = \frac{kQ}{r^2} \quad (9.3)$$

Because from Eq. (9.2), $F = kQq/r^2$, with test charge q, the electrostatic field is

$$E = \frac{F}{q}$$

or

$$F = qE \quad (9.4)$$

The field is therefore the force per unit test charge. *Eq. (9.4) is a definition and true for all electric fields, not just the Coulomb field*. Notice that the dimensions of the E-field are force/charge and the SI units are then newtons per coulomb, $\mathrm{N/C}$. No one's name has gotten attached to the electric field.

Example 9.3: The electric force is large. How many electrons would have to be removed from your body and placed at the center of the earth so that the force between you and the removed electrons would equal your weight? What is the strength of the resulting electric field at the surface of the earth?

Assuming your mass is 50 kg, then with $g = 10\ \text{m/s}^2$, your weight is approximately 500 N. By Coulomb's law we require $500 = kq^2/r^2$, where q is the charge removed from your body. The radius of the earth is approximately 6.4×10^6 m. Plugging in the numbers, $500 = (9 \times 10^9)q^2/(6.4 \times 10^6)^2$. Solving for q gives roughly 1500 C. From Eq. (9.1) this corresponds to about 9.4×10^{21} electrons. By definition, we have $E = F/q = 500\ \text{N}/1500\ \text{C}$, or $1/3\ \text{N/C}$.

Exercise: Estimate the number of electrons in your body and show that the above number represents only one millionth of them. Given that your weight results from the gravitational force between *all* the particles in your body and the enormous number of particles in the earth, approximately 3.5×10^{51} protons and neutrons, this example gives you an idea of how overwhelmingly larger the Coulomb force is compared to the force of gravity.

It is helpful, almost essential to visualize any field as composed of **lines of force** or **field lines**, which are the lines that "conduct" the force. A sphere or point charge radiates field lines uniformly in all directions, as shown in Figures 9.3a and 9.3b. By convention a positive source Q radiates lines outwards, as in (a); field lines converge on a negative source $-Q$ as in (b). Therefore, in (a) a positive test charge q moves radially outwards along the field lines, consistent with the fact that it is repelled by Q, but in (b) moves in radially towards $-Q$.

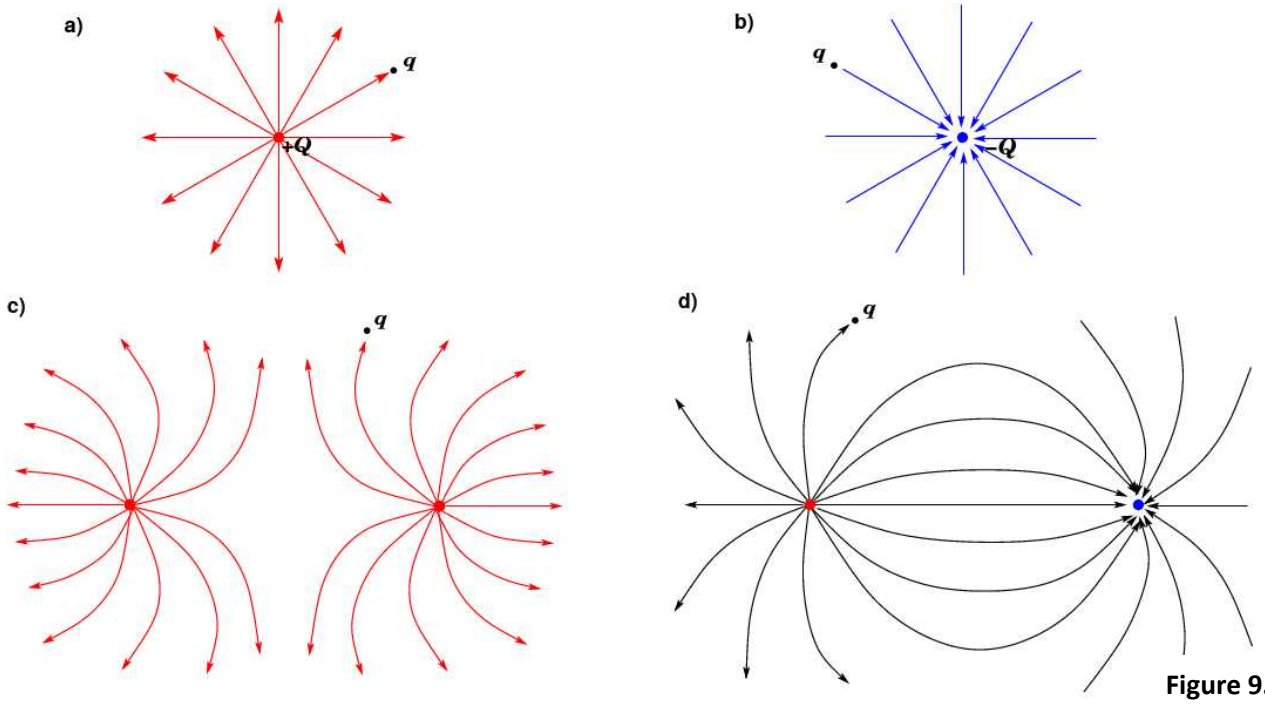


Figure 9.3

Figures 9.3c and 9.3d illustrate the electric fields that result when two point charges are placed near each other. Such fields are termed **dipole** fields, for "two poles." In (c) the dipole consists of two positive charges. Because both source charges repel a test charge, the test charge cannot leave one source charge and reach the other. Instead, it travels along a field line that takes it as far away as possible from both. Dipoles as in (c) are not often found in nature, since the source charges repel each other and the system tends to break apart. In real physical systems, such as molecules, one more frequently encounters dipoles consisting of a positive and negative charge, as in (d). In this case, the test charge is repelled by the positive source charge and attracted to the negative source. Thus, it travels along a field line that begins at the positive source charge and ends on the negative charge.

Any charged object produces an electric field. An infinitely long and wide sheet produces a field, shown in Figure 9.4. For the positively charged sheet, the field lines radiate outward, perpendicularly to the surface. For the negatively charged sheet, the field lines point inward.

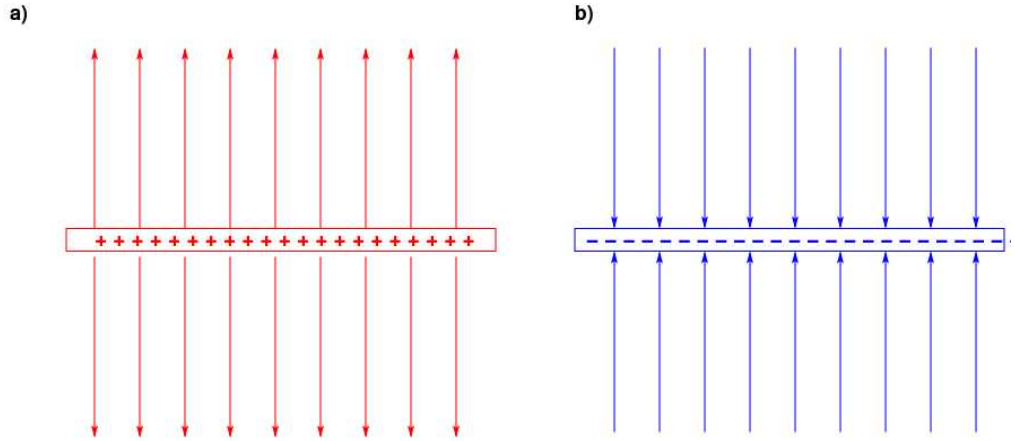


Figure 9.4

One can show that **the electric field E produced by an infinite sheet of charge is constant**. The E-field produced by a point charge or sphere diminishes with the square of the distance. The E-field of a plane does *not*. It is constant. Its exact strength is

$$E = 2\pi k\sigma \quad (9.5)$$

where k is the Coulomb's law constant and σ is the **surface charge density**: the charge per unit area of the sheet in coulombs per square meter. Regardless of how high a test charge q is above the sheets in Figure 9.4, it experiences the E-field given by Eq. (9.5). The force is $F = qE$, and hence also unchanging.[2] *In practice, this means that the dimensions of the sheet should be much greater than the height of the test charge above it.*

Advice: Avoid the common mistake of assuming that all *E*-fields are the fields of point particles.

A common electrical device known as a **capacitor** (next chapter) consists of two oppositely charged plates, one placed above the other. In that case, the field between the plates is twice the value given by Eq. (9.5). You can see why by placing sheet (a) in Figure 9.4 directly above sheet (b).

Note: the combination $4\pi k$ is often written as $1/\epsilon_o$, where $\epsilon_o = 8.85 \times 10^{-12}$ MKS units and is called "the permittivity of free space," an obsolete term from the nineteenth century. Thus, Eq. (9.5) can be rewritten as $E = \sigma/2\epsilon_o$. It makes absolutely no difference which you use and any text that says otherwise is lying.

Exercise: What are the MKS units of ϵ_o?

Example 9.4: If a very long and wide sheet carries a surface charge density $\sigma = 1\ \text{nC/m}^2$, what is the electric field 1 meter above the surface? What is the force on an electron?

As long as the length and width of the sheet are much greater than 1 m, the field can be considered constant and given by Eq. (9.5). Plugging in the numbers gives $E = 2\pi(9 \times 10^9)10^{-9} = 56.5\ \text{N/C}$. The force on an electron would be $F = eE = (1.6 \times 10^{-19})(56.5) = 9 \times 10^{-18}$ N.

---

[2] Roughly, if one views the sheet as composed of an infinite number of point charges, as a test charge moves higher above the sheet, it "senses" the field from ever greater numbers of charges. But these new charges are at ever greater distances and their fields diminish with the square of their distance. The effects of ever more charges but ever weaker fields cancel out to make E constant above the sheet.

Conductors and their fields are particularly important in the study of electricity. As discussed in Section 3, a conductor is any material that allows the free movement of electrons. Electric fields produce forces that push charges. Hence, if a conductor is in equilibrium, meaning that no charges are moving, there can be no electric field. **The electric field on the inside of a conductor in equilibrium is zero**. It can also be shown that in equilibrium **any charges carried by the conductor reside on the surface**. However, any fields parallel to the surface would push the charges around, contradicting the assumption of equilibrium. Therefore, **any electric fields at the surface of a conductor must be perpendicular to the surface**. (The charges cannot be pushed off.) The field at the surface of a conductor thus resembles those illustrated in Figure 9.4.

In this section and the previous we have been concerned primarily with the magnitude of the electrostatic force and electrostatic field. However, *because these quantities have directions, they are vectors*. The SAT subject test expects you only to have a qualitative understanding of the magnitude and direction of the E-field, as in Example 9.4 but does not expect you to handle its vector nature quantitatively; the AP test does. The next section shows how to treat the E-field as a vector. Those interested only in the subject test can skip it.

6. Electric Forces and Fields as Vectors

Eq. (9.2) gave the magnitude of the electrostatic, or Coulomb, force. In most circumstances above the SAT subject-test level it is necessary to treat the electric force as a vector. Coulomb's law should properly be written

$$\mathbf{F}_{12} = \left(\frac{kq_1q_2}{r^2}\right)\hat{\mathbf{r}} \quad (9.6)$$

where $\mathbf{F}_{12}$ is the force that charge 1 exerts on charge 2 and $\hat{\mathbf{r}}$ is the unit vector pointing from charge 1 to charge 2 if the force is repulsive. If the force is attractive, we should write $-\hat{\mathbf{r}}$ instead of $\hat{\mathbf{r}}$. By Newton III we always have $\mathbf{F}_{21} = -\mathbf{F}_{12}$.

Example 9.5: Charges on a triangle. Three charges, $q_1 = q$; $q_2 = 2q$ and $q_3 = 3q$ are arranged on an equilateral triangle, as shown in Figure 9.5a, below. Let $q = 1\,\mu$C and let the sides of the triangle equal 2 meters. What is the total force exerted on q_1 (magnitude and direction) by charges q_2 and q_3?

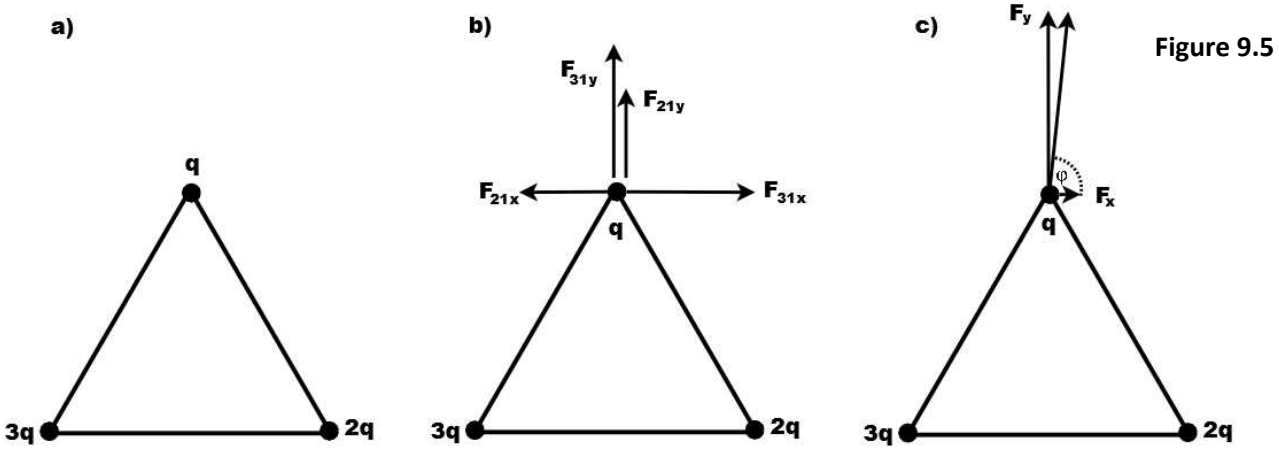


Figure 9.5

Because force is a vector, we need to add the x- and y-components of the forces acting on q_1 separately. By Coulomb's law the magnitude of the force that q_3 exerts on q_1 is

$$F_{31} = kq_1q_3/r^2 = 9\times10^9(10^{-6})(3\times10^{-6})/2^2 = 27/4\times10^{-3}\text{ N}.$$

The x-component of F_{31} is $F_{31x} = F_{31}\cos 60° = F_{31}/2$, or $\frac{27}{8}\times 10^{-3}$N. Similarly, we have $F_{21} = \frac{kq_1q_2}{r^2} = \frac{9}{2}\times 10^{-3}$ N and $F_{21x} = -F_{21}\cos 60° = -9/4\times 10^{-3}$ N, where the minus sign indicates that F_{21x} is pointed in the negative direction, as shown in Figure 9.5b. The total x-component is therefore $F_{1x} = +9/8\times 10^{-3}$ N.

For the y-components, we have $F_{31y} = F_{31}\sin 60° = \sqrt{3}\frac{27}{8}\times 10^{-3}$N, and $F_{21y} = \sqrt{3}\frac{18}{8}\times 10^{-3}$N, both acting up. The total is therefore $F_{1y} = \sqrt{3}\frac{45}{8}\times 10^{-3}$N.

The magnitude of the total force is $F_1 = \sqrt{F_{1x}^2 + F_{1y}^2} \approx 9.8\times 10^{-3}$N. (Reality check: Note that this is almost the y-component, because the x-components have nearly canceled.)

One can express the direction of $\mathbf{F}_1$ one of two ways: We can use unit vectors and write $\mathbf{F}_1 = \left(\frac{9}{8}\,\hat{\mathbf{i}} + \sqrt{3}\frac{45}{8}\,\hat{\mathbf{j}}\right)\times 10^{-3}$N. Alternatively, we can say that the magnitude of $\mathbf{F}_1$ is $F_1 = 9.8\times 10^{-3}$ N and the direction is given by an angle φ such that $\arctan\varphi = F_{1y}/F_{1x}$ above the x-axis, as shown in Figure 9.5c. Plugging in the values for F_x and F_y gives $\arctan\varphi = 5\sqrt{3}$, or $\varphi \approx 83.4°$. Note that this is almost vertical, because the y-component is so much larger than the x-component.

Exercise: Calculate the magnitude and direction of the forces on the other two charges.

Just as the electric force is a vector, **the electric field E is a vector**: $\mathbf{E} = \mathbf{F}/q$. The E-field therefore points in the same direction as the force on the test charge q. One adds E-fields in the same way as one adds the forces in Example 9.5: component-by-component. The total magnitude is then given by the Pythagorean theorem. One *cannot* in general assume that the total magnitude of the field produced by two particles is $E = kQ_1/r_1^2 + kQ_2/r_2^2$, unless the problem is 1-dimensional.

The fact that the E-field is a vector allows one to understand why the field lines must be perpendicular to the infinite planes in Figure 9.4. Let each $(+)$ drawn in the plane represent a point charge q. Its field lines thus extend radially outward and have components in every direction: to the left and right $(\pm x)$, in and out of the page $(\pm y)$, up and down $(\pm z)$. Pick any q to be the "center." Because the plane is infinite, for every q to the left of your center, there will be a q an equal distance to the right with the x-component of its field pointed in the opposite direction. Thus, all the x-components cancel. All the y-components cancel for the same reason, leaving only the z-components (up and down), which add. This is an example of the use of **symmetry** to arrive at an answer before doing any calculations. **Develop your sense of symmetry**.

7. Electric Current

Just as the motion of water molecules produces a water current, the motion of electric charges produces an **electric current**. Any current is defined as the amount of something that flows past an observation point per unit time. Electric current, almost always designated I, is defined simply as the amount of charge flowing past a given point per second:

$$I = \frac{\Delta Q}{\Delta t} \qquad (9.7)$$

Because Q is measured in coulombs and t in seconds, the dimensions of current are charge/time and the SI unit of current is coulombs per second, which has been given the name ampere, after André Marie Ampère: $1\,\text{A} \equiv 1\text{C}/1\text{s}$. Because a coulomb is a large unit of charge, the amp is a rather large unit of current. Many modern devices run on currents of milliamps, mA.

Although in real conductors it is the negatively charged electrons that actually move, in physics, by convention, **currents are taken to be positive charges**. If in a given situation the electrons are moving one way, the "conventional" current is therefore moving in the opposite direction. This sign convention is stupid, but becomes important in getting the direction of certain quantities correct, like the direction of magnetic forces. Because currents have a direction, they can be considered vectors and some texts write **I** when direction becomes important. Electrical currents can flow only in complete **circuits**. A circuit is any closed loop that conducts electricity. It may be as simple as a wire loop or contain other devices such as **resistors** and **capacitors**. In this chapter, any circuit mentioned is merely a wire loop or coil. We discuss resistors and capacitors in Chapter 10.

Example 9.6: A lot of current? An Avogadro's number of charges passing past an observation point per second represents how large a current?

Avogadro's number is 6.02×10^{23}. A coulomb of charge is 6.24×10^{18} electronic charges. Therefore, Avogadro's number represents slightly less than 10^5 coulombs and an Avogadro's number of charges per second is nearly 10^5 amps. Such large currents are not generally encountered in daily life; household circuit breakers trip at 15 or 20 amps. However, if a lightning bolt transfers 10 C of charge in a millisecond, that represents 10^4 amps, and currents of over one million amps (mega amps, MA) are routinely employed in the experimental nuclear fusion devices called tokamaks.

8. Magnetic Fields

The ancients knew that certain materials, specifically lodestone, attracted iron and could also be used as compasses. The word **magnet** is in fact Greek for "stone from Magnesia," the region in Turkey where lodestone was commonly found. Today magnets are a common feature of everyday life; however, their description in terms of physics is complicated.

A permanent magnet is one that, as its name implies, produces its own magnetic field. Permanent magnets attract **ferromagnetic materials**, in particular iron and nickel. All magnets have a **north pole** (N) at one end and a **south pole** (S) at the other. Similarly to electric charges, **like magnet poles repel and unlike poles attract**. Thus, the north poles or south poles of two nearby magnets repel each other, while a north and a south pole of two magnets attract. Like the signs for electric charges, the designation for magnetic poles is arbitrary, but by convention the north pole of a compass needle is the end that points to the earth's north magnetic pole in the Arctic. However, since opposite poles attract, that means that what is called the earth's north magnetic pole is really its south magnetic pole.

There is, however, a fundamental difference between magnetic poles and electric charges: If you break a magnet in half, the result is *not* an isolated south and a north pole; the result is a new magnet with a north and a south pole. You can continue this process indefinitely but the result will always be a new magnet with a north and south pole.

Isolated magnetic poles do not exist

Isolated magnetic poles are termed "monopoles," but none have ever been found. As far as anyone can tell, the nonexistence of monopoles is a law of nature.

Because all magnets come with a north and south pole, magnets do not produce anything that resembles the electric field of a single charge discussed in Section 5. **The field of a magnet is a dipole field**. Thus, it exactly resembles the field in Figure 9.3d. The positive source charge merely needs to be replaced by a north pole and the negative charge by a south pole, as in Figure 9.6. By convention, **magnetic field lines leave the north pole and arrive at the south pole**. If you have ever sprinkled iron filings on a paper above a bar magnet, you have seen this field with your naked eye.

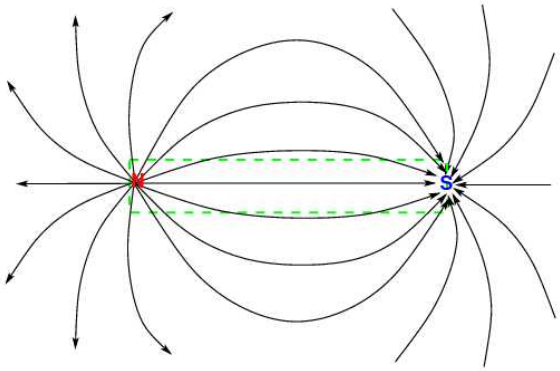


Figure 9.6

For mysterious reasons, magnetic fields are generally designated by the letter B. Because B-fields are dipolar, their mathematical description is much more complicated than for the electrostatic field of a single charge; there is no simple "Coulomb's law" for the field.[3] You will not be required to write down the strength of a dipole field on any SAT exam, but you will be required to recognize a dipole field when you see one.

9. The Magnetic Force

How does one determine B? Like electric fields, magnetic fields exert forces on charged particles. The strength of the B-field is in fact defined by the magnitude of the force it exerts on an electric charge. Unfortunately, the magnetic force bears *no* resemblance to Coulomb's law. In understanding the force that a magnetic field exerts on a charge, it is essential to keep in mind that **B** has a direction; it is a *vector*. The *magnitude* that the force B exerts on q is given by the

Lorentz Force Law

$$F = qvB\sin\theta \quad (9.8)$$

after Hendrik Lorentz. Like Coulomb's law, the Lorentz force law may be taken as a law of nature. In Eq. (9.8), q is the test charge, the charge *on which* the force is acting, v is the test charge's speed, B is the strength of an *external* magnetic field through which q is moving and θ is the angle between the charge's velocity and the direction of the field at the position of the particle (Figure 9.7). Notice that B *does not equal* F/q.

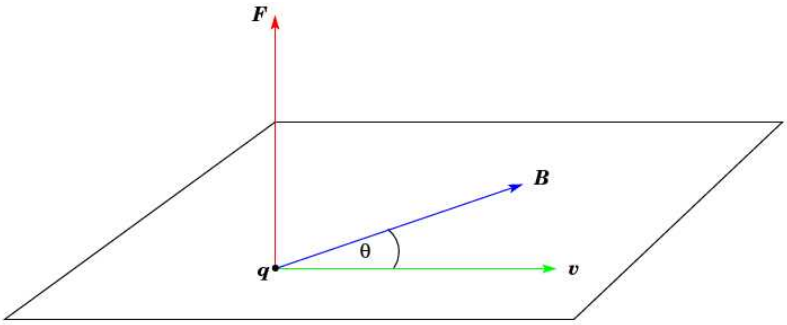


Figure 9.7

The crucial point illustrated by the figure is that **the magnetic force on q acts at right angles to both q's velocity and to the B-field.**

[3] Far away from the dipole, the B-field obeys an "inverse-cube" law; it diminishes with the cube of the distance, but also depends on direction.

To get the direction of F use the right-hand rule (RHR) discussed in Section 1.7: point the fingers of your *right* hand in the direction of **v** and curl them in the direction of **B**; your thumb points in the direction of **F**.

Commentary on the Lorentz force law: Several important features about magnetic forces are immediately apparent from Eq. (9.8):

1) The force depends on the charge's velocity. If $v = 0$, a B-field exerts *no* force on an electric charge.

2) If the charge's velocity **v** is in the *same* direction as **B** then the B-field exerts *no* force on the charge. ($\theta = 0$, $\sin\theta = 0$.)

3) If **v** is at right angles to **B**, then the magnitude of the force is simply $F = qvB$. ($\theta = 90°, \sin\theta = 1.$) *This is the situation encountered in most elementary problems.*

4) The direction of the force given by the RHR assumes that q is positive. A negative charge results in a force in the opposite direction. Using your left hand in finding the direction of F gives the result for a negative charge.

5) Measuring the force on a charge gives the size of B: $B = F/qv$. It also gives the appropriate units for the magnetic field: B should equal $\text{newtons}/(\text{coulombs} \times \text{m/s})$. But 1 coulomb/s is 1 amp, and so we arrive at the units for the B-field: $1 \text{ tesla} \equiv 1 \text{ newton}/(\text{amp meter})$, or $1\text{ T} \equiv 1\text{N}/(\text{A m})$.

Tip: A tesla is an extremely large unit. The strongest magnets in the world, which can levitate frogs (google "levitating frog"), are about 60 T. Those under construction for the world's largest tokamak will be about 15 T. Most magnets have B-fields much smaller than 1 T. The earth's magnetic field is about 0.5×10^{-4} T. For this reason, the cgs unit for magnetism, the gauss, is often preferred. 1 G $\equiv 10^{-4}$ T.

Eq. (9.8) gives the force on a single moving charged particle. We know, however, that moving charges constitute an electric current; a single moving charge is essentially a (small) electric current. Usually currents are found in wires. For a wire of length L carrying a current I, Eq. (9.8) can be rewritten as

Lorentz Force on a Wire

$$F = ILB\sin\theta \qquad (9.9)$$

Here, θ has become the angle between the direction of the current **I** and the magnetic field **B**. (We have written the current as a vector **I** because the direction is now important. Alternatively, we can write the wire's length as a vector, **L**, whose direction is in the direction of the current. The two ways are equivalent.)

It is not difficult to get from Eq. (9.8) to Eq. (9.9). If we imagine a segment of a wire carries N charges, then the total Lorentz force on that segment should be simply $F = NqvB\sin\theta$. But Nq is the total charge Q and dimensionally $v = \text{length/time}$. Thus, Nqv must be LQ/t and $Q/t = I$. Then $F = ILB\sin\theta$.

More precisely, a cylindrical wire segment with length L and cross-sectional area A has volume $V = LA$. If the density of charges is n particles per unit volume, then $N = nV = nLA$, and so $F = LqAnvB\sin\theta$. However, in Section 3.14 we showed that Anv is the number of particles passing through the area A per unit time, and therefore $qAnv$ is the total charge passing through A per unit time. But this is the definition of a current, I. Consequently, $F = ILB\sin\theta$.

Example 9.7: Current loop in B-Field. Suppose a current I of 1 amp flows in a counterclockwise direction around a triangular loop with sides 3 m, 4 m and 5 m, as shown in Figure 9.8. A B-field of 0.2 T is pointing into the page. (The symbol ×, indicating an arrow tail, is universally used to indicate a vector pointing into the page; whereas the arrowhead, ●, universally indicates a vector pointing out of the page.) What are the direction and magnitudes of the forces on each segment?

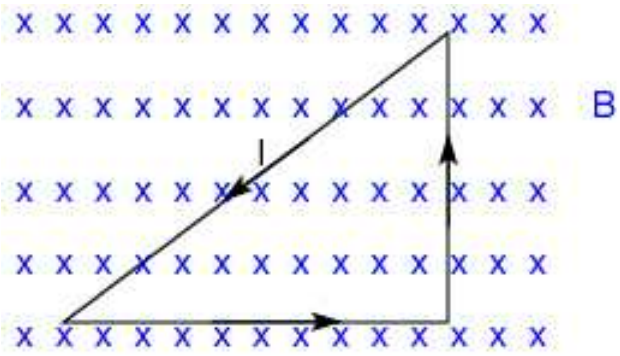


Figure 9.8

Notice first that the angle between the magnetic field and all the wire segments is 90°, and so by Eq. (9.9), the force on the vertical segment is $F = ILB = 1 \times 3 \times 0.2 = 0.6$ N. The RHR applied to this segment shows the force is acting directly to the left. The force on the horizontal segment is $F = 1 \times 4 \times 0.2 = 0.8$ N and by the RHR is directed straight up. The force on the hypotenuse is $F = 1 \times 5 \times 0.2 = 1$ N, directed into the triangle perpendicularly to the hypotenuse. (**Exercise:** draw in the force vectors.)

Note that the total force acting on the loop must be zero. (Why?) Thus, the magnitude of the force acting on the hypotenuse is equal to the magnitude of the forces acting on the other two sides, and all the x- and y- components of the force vectors must cancel out. (**Exercise:** Show this explicitly by finding the components of the force vectors you have drawn.)

A further important property of the magnetic field is demonstrated by returning to the force exerted on a single charge q:

Example 9.8: Motion of charge in B−field. Suppose a charge q is moving to the left with velocity **v** along the $-y$-axis, and that a uniform **B** field points up along the positive z-axis. In what direction is **F**? What path does q follow?

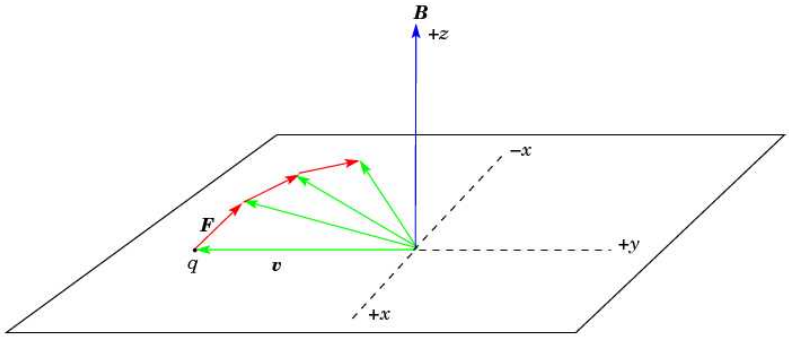


Figure 9.9

Since the problem states that **B** is uniform, it is the same everywhere and although in Figure 9.9 we have drawn only one arrow, you are to imagine **B** pointing upward everywhere on the plane. By the RHR, the force **B** exerts on q is into the page, along the $-x$-axis, at right angles to **B** and **v**.

Because **F** is at right angles to **v**, it changes the direction of **v** into the page, as shown. Now **v** has a slightly different direction, as shown by the second velocity arrow. Applying the RHR to the new **v** shows that the force is now pointed in a slightly different direction, indicated by the second force arrow. Once again, the direction of the velocity changes slightly, to the third velocity arrow. The RHR now gives a new force in the direction of the third force arrow, and so on. **The resulting motion of the charge is a circle around the magnetic field lines**.

Suppose, however, that the particle has a velocity $\mathbf{v} = \mathbf{v}_x + \mathbf{v}_z$. In other words, it now has a velocity component parallel to **B**. Since the angle θ between **B** and $\mathbf{v}_z$ is zero, **B** exerts no force on the particle in that direction and $\mathbf{v}_z$ remains constant. Thus, the particle goes around in a circle in the xy-plane, as drawn, but also moves upward along the z-axis at a constant velocity; it therefore executes a spiral around the field lines.

Further, note in this example that because the force is not accelerating the particle in its direction of motion **the speed of the particle has not changed**. More formally, from Eq. (3.16) the power expended in moving a particle is $P = Fv\cos\varphi$, where φ is the angle between the force **F** and the velocity **v**. In this case the force is acting at right angles to the velocity; $\varphi = 90°$ and $P = 0$. Consequently, the particle's kinetic energy and speed v remain constant. Since there is no change in kinetic energy, one can say by the work-energy theorem

A magnetic field does no work

10. The Magnetic Force as a Vector

This section is at the AP level.

Eq. (9.8) can be written more compactly as the vector equation

$$\mathbf{F} = q\mathbf{v} \times \mathbf{B} \quad (9.10)$$

where $\times$ represents the cross product already encountered in Sections 1.7 and 4.6. Similarly, the force on a wire is given by

$$\mathbf{F} = I\mathbf{L} \times \mathbf{B} \quad (9.11)$$

where **L** is the length of the wire taken as a vector pointed in the direction of the current I.

Exercise: If a current-carrying wire runs $53°$ above the $+x$-axis, and a B-field is pointed in the $-y$ direction, in what direction is the magnetic force on the wire?

11. Electric currents produce magnetic fields

In the previous sections the B-field was assumed to be given, and in Section 8 they were assumed to be produced by permanent magnets. More often encountered in basic physics and on the SAT exams are the B-fields resulting from Oersted's original discovery uniting electricity and magnetism: **electric currents generate magnetic fields**. Indeed, it is these magnetic fields that find their way into countless everyday applications. In elementary courses essentially only two situations arise: the B-field produced by a wire coil and the B-field produced by a long a straight wire.

A coil of wire wrapped into a cylinder is termed a **solenoid** (Figure 9.10). When a current I is passed through the wire it produces a B-field closely resembling the field of a bar magnet; in other words, it creates a dipole field like the one in Figure 9.6. Current-carrying solenoids are the basis of **electromagnets**, among other things.

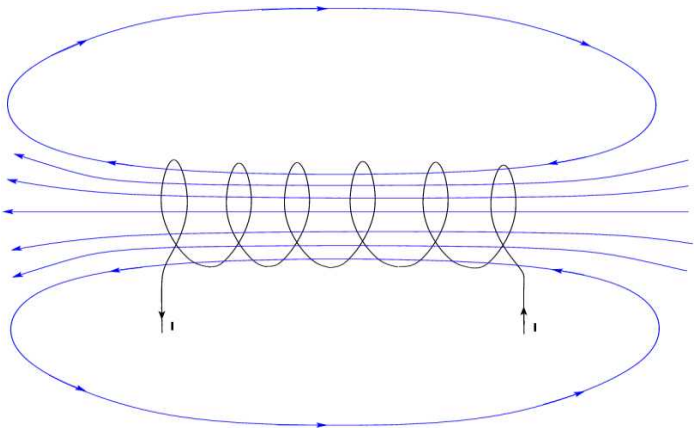


Figure 9.10

If the solenoid is long enough, the B-field *inside* is essentially constant. We state without proof that the magnitude of a solenoid's interior B-field is given by

$$B = \mu_o n I \quad (9.12)$$

where n is the number of turns of the wire per unit length. The message of Eq. (9.12) is that **the magnetic field strength in a solenoid is directly proportional to the size of the current**. The constant of proportionality, μ_o, is called the "permeability of free space" (another obsolete term dating from the nineteenth century) and has the value $\mu_o = 4\pi \times 10^{-7}$ SI units.

Example 9.9: Typical solenoid. If a solenoid is thirty-three centimeters long, has 1000 turns of wire and carries a current of 2 amps, what is its B-field?

If the solenoid is thirty-three centimeters long and carries 1000 turns, then $n = 1000/0.33 = 3000$ turns per meter. Then $B = 4\pi \times 10^{-7} \times 3 \times 10^3 \times 2 = 7.5 \times 10^{-3}$ T.

Exercise: From Eq. (9.12) what are the exact units of μ_o in terms of B, I and n?

The other B −field of interest in elementary physics is the B-field around a long, straight wire. Its magnitude is given by

$$B = \frac{\mu_o I}{2\pi r} \quad (9.13)$$

This equation is known as "Ampère's circuital law," although Ampère never wrote it down. It states that **the strength of the magnetic field around a wire is proportional to the current *I* and inversely proportional to the distance from the wire *r***. *This is not an inverse-square law*.

As shown in Figure 9.11, the current I through the wire produces a B-field that forms concentric circles around the wire, **The direction of the *B*-field is found by curling the fingers of your right hand around the wire such that your thumb points in the direction of the current. The direction of your fingers points in the direction of the *B* −field.** (This is sometimes called right-hand-rule No. 2; Section 4.6).

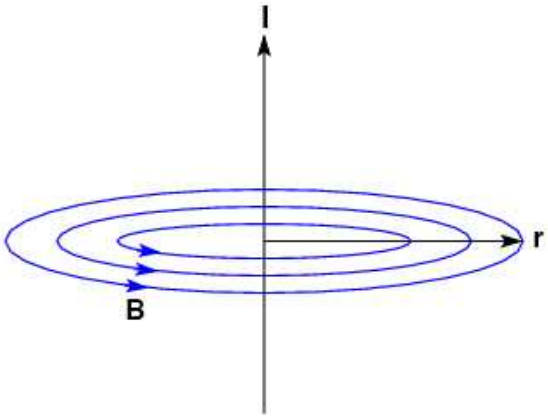


Figure 9.11

12. Electromagnetic Induction, Basic

Oersted discovered that electric currents produce magnetic fields. Under the influence of Newton's third law, action and reaction, natural philosophers of the time were convinced that the reverse should also be true: magnetic fields should produce electric currents. This turned out to be correct—almost.

In 1831 Michael Faraday in England and Joseph Henry in America performed a series of simple experiments. In one, Faraday merely thrust a bar magnet through a coil of wire he had wrapped around a paper tube, as illustrated in Figure 9.12. He discovered that while the magnet was *moving*, the needle on a current meter deflected. When the magnet was *stationary*, *no deflection* of the needle occurred. When the magnet was removed from the tube, the needle momentarily deflected in the opposite direction, and the action of the needle was also reversed when the south pole of the magnet was passed through the tube instead of the north pole.

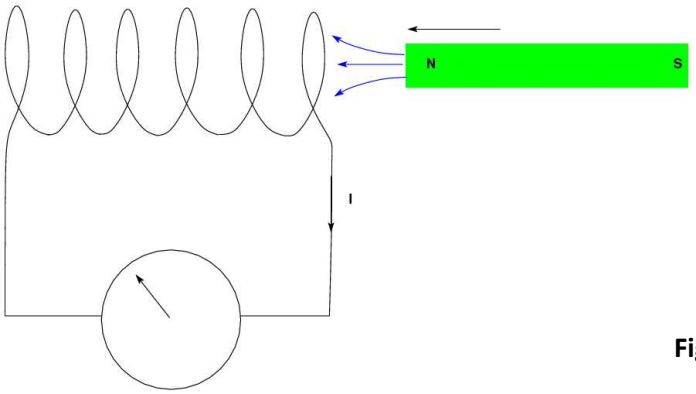


Figure 9.12

We say that the motion of the magnet has *induced* a current in the solenoid. It does not matter whether the magnet is accelerating or not. What is important is that the **magnetic flux** in the coil is *changing* with time. If the B-field is perpendicular to the end of the coil, the magnetic flux is defined by $\Phi = BA$, where A is the area of the coil. Thus, you can change Φ ("Phi") by changing B, A or both. In a coil, A remains constant, so the flux changes when B changes. By moving the magnet in and out of the coil, you are changing the flux and therefore you produce a current.

Thus, Faraday and Henry established that *changing* magnetic fields produce electric currents, which was almost—but not quite—the converse of Oersted's discovery. **Faraday's law of induction**, as it is called, provides not only the principle of operation for cordless tea kettles, but for the electrical transformer and hence the entire power industry. Electromagnetic induction also accounts for the existence of electromagnetic waves and the possibility of radio and television. It can safely be called the basis of modern civilization.

For the subject test, only a qualitative understanding of electromagnetic induction is required. Quantitative details are provided in Section 14, which is at the AP level.

13. Electric Potential

Until now we have been concerned with forces and fields. The fact that according to Coulomb's law two unlike charges attract each other indicates that it would take work, or energy, to move them apart. Indeed, because Coulomb's law is identical in form to Newton's law of gravity, the potential energy associated with two electrical point charges q_1 and q_2 is identical in form to gravitational potential energy already discussed in Sections 3.6 and 3.7

$$U = \frac{kq_1q_2}{r} \quad (9.14)$$

where r is the distance between the charges. This expression also holds for two spherical charge distributions in which case r is the distance between their centers. Eq. (9.14) is referred to as the **Coulomb potential**. Its most important feature is that it diminishes with r and goes to zero as $r \to \infty$.

Apart from the fact that the Coulomb potential involves charges while the gravitational potential involves masses, the main difference between the gravitational and Coulomb potentials is that there is no minus sign in Eq. (9.14), as there was in Eq. (3.8). This is because the gravitational force is always attractive, and gravitational potential energy always negative. However, q_1 and q_2 can be either positive or negative and so the potential can be either as well, as seen in Figure 9.13.

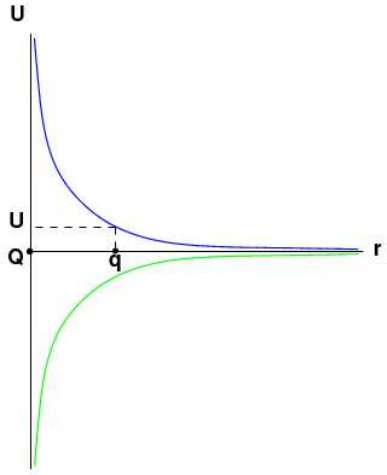


Figure 9.13

In this figure we have put a source charge Q at the origin and a test charge q at a distance r from Q. Assuming that both charges are the same sign, the potential energy curve is the top one. (If Q and q were of opposite signs, the potential energy curve would be the bottom one.) The potential energy U of q is marked. It is important to note that although r is the distance between the charges, the potential energy U is the energy above zero, *which is its value at* $r = \infty$. **All energies are energy differences**. You can set the zero point where you like. **As with gravity, electrical potential energy is measured with respect to the potential at infinity.**

Because potentials are scalars, it is generally easier to work with them than with forces or fields, which are vectors. In particular,

Potentials add

For example, if we put two source charges, Q_1 and Q_2, at the origin, then q's total potential energy is simply $U = kQ_1q/r + kQ_2q/r$. If this is not obvious, the best way to see it is to recall that

The electric field is a conservative field; conservation of energy holds; the work-energy theorem holds.

In other words, source charge Q_1 repels q. Thus, it requires work to bring q from infinity to a final position a distance r from Q_1. The amount of work is exactly the potential energy with respect to infinity: $U_1 = kQ_1q/r$. (If you have forgotten the work-energy theorem, see Section 3.9.) At the same time, Q_2 is also repelling q and it requires work $U_2 = kQ_2q/r$ to bring in q against the repulsive force of Q_2. Thus, the total potential energy is $U_1 + U_2$.

If we have three particles, as in Figure 9.5a, the situation is the same. The total potential is the energy needed to bring in each particle from infinity, assuming the others are acting as source charges already in place: $U = kq_1q_2/r_{12} + kq_1q_3/r_{13} + kq_2q_3/r_{23}$, where r_{12} is the final distance between q_1 and q_2, etc.

Example 9.10: Flying charges. Suppose that the charge q_1 in Example 9.5 has a mass $m = 10^{-9}$ kg. At a time $t = 0$ it is suddenly allowed to fly off from the triangle. What is its final speed at a large distance from the triangle?

The initial potential energy of q_1 in the triangle is

$$U = \frac{kq_1q_2}{r_{12}} + \frac{kq_1q_3}{r_{13}} = \frac{9 \times 10^9}{2}(1 \times 2 + 1 \times 3) \times 10^{-12} = \frac{45}{2} \times 10^{-3}\ \text{J}.$$

By conservation of energy, this potential must equal the final kinetic energy of q when $r \to \infty$ and the potential becomes zero. Thus, we equate the above to $1/2\, mv^2$. With $m = 10^{-9}$ kg, we get $v^2 = 45 \times 10^6\ \text{m}^2/\text{s}^2$, or $v = 6.7 \times 10^3\ \text{m/s}$.

If you are wondering why we didn't use the *total* potential energy of the triangle ($U = kq_1q_2/r_{12} + kq_1q_3/r_{13} + kq_2q_3/r_{23}$), imagine bringing in q_1 to the triangle from infinity. The work required is the work done against q_2 and q_3, as written in the solution. This is the amount of energy released when q_1 flies off. In other words, q_2 and q_3 remain in place, so the potential energy between them is not used.

Just as we defined the *E*-field as force per unit test charge, it is extremely useful to define potential energy per unit test charge. If $U = Qq/r$, then the potential energy per unit test charge is

$$V = \frac{U}{q} \qquad (9.15)$$

The quantity V is referred to as electric potential (without the "energy") but is more commonly known as **voltage**. The dimensions of voltage are energy/charge. One **volt** is defined as one joule per coulomb: $1\text{V} \equiv 1\ \text{J/C}$. A joule, however, is a newton meter, and so $1\ \text{V} \equiv \text{N m/C}$.

Warning: Physicists and engineers often use the terms "potential" and "potential energy" interchangeably. Remember that electric potential (voltage) has different dimensions and units than potential energy.

Because all energies are actually energy differences, Eq. (9.15) should be written $\Delta U = q\Delta V$. Usually when voltage is given in a problem it is understood that it is a voltage difference with respect to some reference point.

Further, we already know that $F = qE$. We also know that for a *constant* force, work is the force multiplied by the distance over which the force is applied, $W = Fd$, and according to the work-energy theorem $W = -\Delta U$. Thus, with $F = qE$,

$$\Delta U = -qEd \quad (9.16a)$$
or
$$\Delta V = -Ed \quad (9.16b)$$

Because these are just potential differences, unless we are really concerned whether they are plus or minus, we usually just write $U = qEd$ and $V = Ed$. Although we have defined E as newtons per coulomb, note that this last expression gives the more common unit for the electric field: $[E] = \mathrm{V/m}$, volts per meter. **Eqs. (9.16) apply only to constant fields.**

We will make ample use of voltage in the next chapter when we review simple circuits. For now, we note that the expression $U = qV$ suggests a convenient unit of measurement. The **electron volt** (eV) is defined as the energy acquired by an electron when it passes through a potential difference of one volt. Since for an electron $q = e = 1.6 \times 10^{-19}$ C, one electron volt equals 1.6×10^{-19} C $\times$ 1 V, or

$$1 \text{ eV} = 1.6 \times 10^{-19} \text{ J} \quad (9.17)$$

The eV is the standard unit of measurement for energy in atomic physics (Chapter 12), when one deals with charges on the order of one electronic charge and energies on the order of 1 eV. In general, it is a good idea to use units that are about the same size as the quantities in the system under measurement.

14. Electromagnetic Induction, Quantitative

To understand how electromagnetic induction, qualitatively described in Section 9.12, comes about, one must define **magnetic flux**. In general, flux is defined as the flow of some quantity that passes through a surface. The water flux from a hose sprayed through a tennis racket will have its maximum value if the tennis racket is oriented perpendicularly to the stream. The flux through the racket will be zero if it is oriented parallel to the stream.

The situation for the magnetic field is essentially the same. A reasonable definition of magnetic flux would be the total number of B-field lines going through a surface. It is difficult, however, to count field lines. Instead, the magnetic flux Φ passing through a surface of area A is defined to be

$$\Phi = BA\cos\theta = \mathbf{B} \cdot \mathbf{A} \quad (9.18)$$

where θ is the angle between the normal to the surface **n** and the magnetic field **B.** The meaning of the definition is illustrated in Figure 9.14.

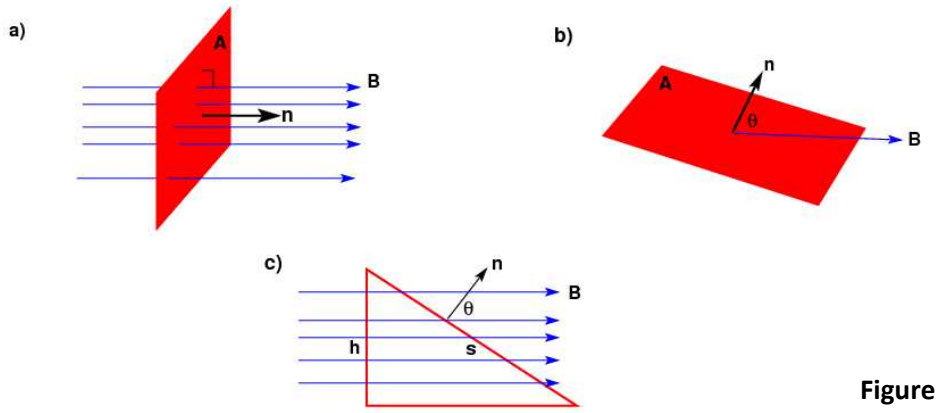


Figure 9.14

If **B** is perpendicular to the surface A, as in (a), then it is parallel to the normal vector **n** and $\theta = 0$. In that case, $\Phi = BA$. *This is the situation encountered in most elementary problems*. However, like the tennis racket, the surface may be tilted at an angle θ with respect to **B**, as in (b). If $\theta = 90°$, then no magnetic field lines pass through the surface and $\Phi = 0$. Figure (c) shows a side view of the B-field passing through the surface at an angle. The full area A equals the hypotenuse s times the width of the surface (into the page, not shown). However, the same number of field lines pass through the vertical face (height h) as through the slant face. Thus, if flux is to measure the number of field lines passing through a surface, the flux through both faces should be the same. From the diagram, it is easy to see that $h = s\cos\theta$. (**Exercise**: See this.) Thus, the area of the vertical face, which is perpendicular to the B-field, is $A_\perp = A\cos\theta$. The flux through both faces is $\Phi = BA_\perp = BA\cos\theta$. If one defines area as the vector $\mathbf{A} = \mathbf{n}A$, then one can write the flux as $\Phi = \mathbf{B}\cdot\mathbf{A}$, as we did above.

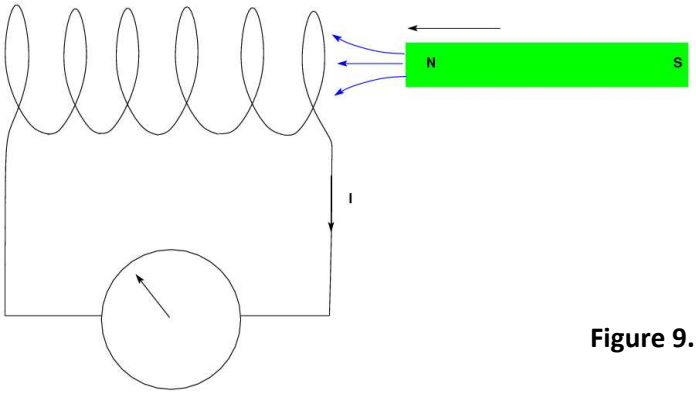


Figure 9.15

Faraday's results have already been described in Section 9.12: a moving bar magnet induces a current in a solenoid (as shown in Figure 9.10, and reproduced here). In modern terms **Faraday's law of electromagnetic induction** can be expressed as

$$\mathcal{E} = N\frac{\Delta\Phi}{\Delta t} \qquad (9.19)$$

The script $\mathcal{E}$ is used by most texts to denote "electromotive force," or EMF, but it is just the voltage induced in the circuit by the motion of the magnet. N is the number of turns in the coil. The meaning of Faraday's law is: **the voltage induced in a circuit is equal to the rate of change of the magnetic flux through the circuit**.

Commentary on the law of induction:

1) Steady currents produce B-fields, as in Eqs. (9.12) and (9.13), but according to Faraday's law, only a *changing* magnetic flux can produce voltages in circuits. Voltages in turn produce currents (next chapter). Therefore, Henry and Faraday did not quite establish that B-fields produce electric currents.

2) The flux threading the coil is BA. For there to be a nonzero voltage $\mathcal{E}$ and current I induced into the circuit, B or A must therefore change with time, or both.

3) In more advanced texts a confusing minus sign $(-)$ appears before the right hand side of Eq. (9.19). Don't worry about it.

4) If necessary, the direction of the current induced in the coil can be found by using **Lenz's law**, which states that **the direction of the induced current must be such as to oppose the change in flux**. In the diagram, the flux is *growing* because the motion of the bar magnet toward the coil increases the B-field inside the coil. The induced current I must therefore be down the outside of the coil, as drawn. Why? Align your right thumb with the direction of I. According to RHR No. 2, your fingers are curling in the direction of the induced B-field, which opposes the one drawn. In other words, the coil acts like a magnet with a north pole that repels the north pole of the bar magnet. This is necessary because of conservation of energy. If it were the other way around, the induced B would attract the bar magnet, pulling it faster into the coil, in turn generating a greater flux, in turn pulling the magnet even faster…Perpetual motion machines are a violation of conservation of energy.

5) Notice that the coil in this circuit is the same as the solenoids already discussed in Section 11.

6) The dimensions of flux are BA and the SI unit of flux is the weber. $1\,\text{weber} \equiv 1\,\text{tesla} \times 1\,\text{m}^2$; or $1\,\text{Wb} \equiv 1\,\text{T}\,\text{m}^2$.

Example 9.11: Generating currents. A 0.1 T bar magnet is initially far from a solenoid with 1000 turns and cross sectional area of 75 cm^2. You thrust the magnet into the coil in a time 0.5 s. What is the voltage induced in the solenoid?

We assume that the B-field in the coil is initially zero and rises to 0.1 T in 0.5 seconds. Since the area A of the coil is constant, $\Delta\Phi/\Delta t = A\Delta B/\Delta t = 7.5 \times 10^{-3} \times .2 = 1.5 \times 10^{-3}\ \text{Wb/s}$. By Eq. (9.19) the induced voltage is therefore 1.5 V, enough to run a flashlight, briefly.

Example 9.12: Motional EMF. An important application of Faraday's law is called "motional EMF" (for electromotive force) and provides the basis for the rail gun, currently under development by the Navy (Many online videos and Problem 9.50. A constant, uniform B-field is directed into the page. You pull a conducting bar to the right across conducting rails with a constant velocity **v**, as shown in Figure 9.15. What is the voltage induced in the circuit composed of the bar and rails in terms of v, B and other quantities you might need? What direction is the induced current?

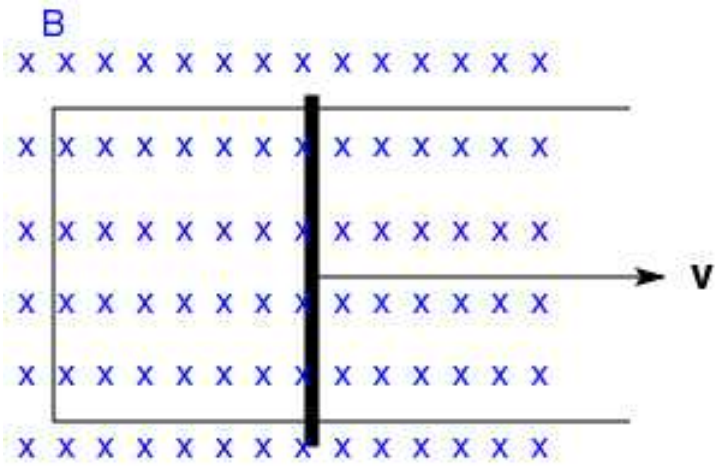


Figure 9.16

Call the length of the bar ℓ and let x be the distance from the left end of the rail to the bar. The area of the region enclosed by the bar and rails is $A = \ell x$. By definition, the flux passing through this region is $\Phi = BA$. As you pull the bar to the right you are increasing x and hence A. Because B and ℓ are constant, the change in flux is $\Delta\Phi/\Delta t = B\ell\Delta x/\Delta t$. But $\Delta x/\Delta t = v$, and so $\Delta\Phi/\Delta t = B\ell v$, which by Faraday's law is also the voltage $\mathcal{E}$ induced into the circuit: $\mathcal{E} = B\ell v$.

Because free electrical charges in the conducting bar are moving to the right, they will experience a force given by the Lorentz force law. The RHR shows the direction of the force is up; consequently, the direction of the current is also up and proceeds counterclockwise around the rails. (Draw this in.) Notice that this induced current I, by RHR No. 2, produces its own B-field *out* of the page, which lessens the total B-field and consequently opposes the increase in flux, in accordance with Lenz's law.

If $B = 1$ T, $\ell = 0.2$ m and $v = 100$ m/s, then the induced voltage is 20 V. The size of the induced current will depend on the resistance of the bar (next chapter). This example provides a simple case of an electromagnetic **generator**, any device that transforms motion into electric currents.

ELECTRICITY AND MAGNETISM PROBLEMS

Problems 1-20 are at the subject-test level; the remainder are at the AP level.

1. Two identical conducting spheres, A and B carry charges $+Q$ and $-Q$, respectively. A third, identical conducting sphere C carries charge $Q = 0$. Sphere A is touched to sphere C and separated. Next, sphere B is touched to sphere C and separated. Finally, A is touched to B and separated. What is the final charge on each sphere?

   (A) $A = Q\ ;\ B = -Q;\ C = 0$
   (B) $A = Q/2\ ;\ B = Q/2\ ;\ C = Q/4$
   (C) $A = Q/8\ ;\ B = Q/8\ ;\ C = -Q/4$
   (D) $A = Q/2\ ;\ B = -Q/4\ ;\ C = -Q/4$
   (E) $A = Q/4\ ;\ B = -Q/8\ ;\ C = -Q/4$

Questions 2 - 3 refer to the following figure.

2. Two conducting spheres of differing radii are in contact, as shown. A positively charged rod is touched to the large sphere and removed to a great distance. The large and small spheres are then separated. We can then say that

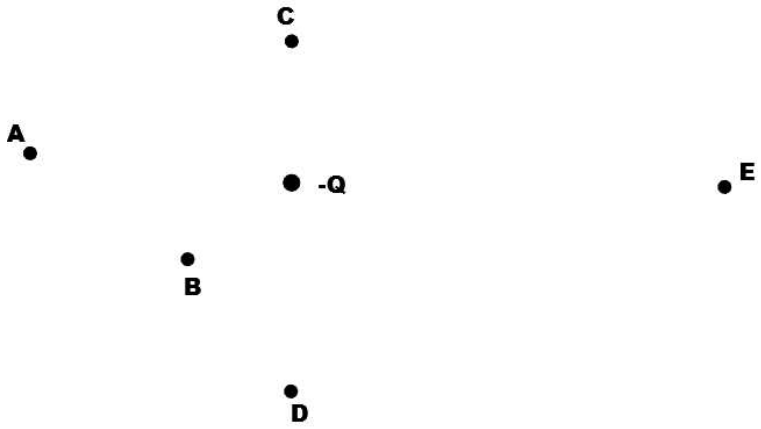
   (A) the big sphere will be positively charged and the smaller sphere will be negatively charged.
   (B) both spheres will be positively charged.
   (C) the big sphere will be negatively charged and the smaller sphere will be positively charged.
   (D) both spheres will be negatively charged.
   (E) all the charge will migrate to the smaller sphere.

3. The positively charged rod is brought near the large sphere, but without touching it. The two spheres are separated and lastly the rod removed to a distance. We can then say that

   (A) the big sphere will be positively charged and the smaller sphere will be negatively charged.
   (B) both spheres will be positively charged.
   (C) the big sphere will be negatively charged and the smaller sphere will be positively charged.
   (D) both spheres will be negatively charged.
   (E) all the charge will migrate to the smaller sphere.

Questions 4 – 6 refer to the following information.

All the charges in the diagram below are negative.



4. A negative charge $-Q$ is located at the position shown. At which of the labeled points is the electric field smallest?

   (A) A
   (B) B
   (C) C
   (D) D
   (E) E

5. At which pair of points is the field most nearly the same?

   (A) B and C
   (B) B and A
   (C) D and E
   (D) C and D
   (E) A and D

6. At which point will a negative charge experience a force directly downward?

   (A) A
   (B) B
   (C) C
   (D) D
   (E) E

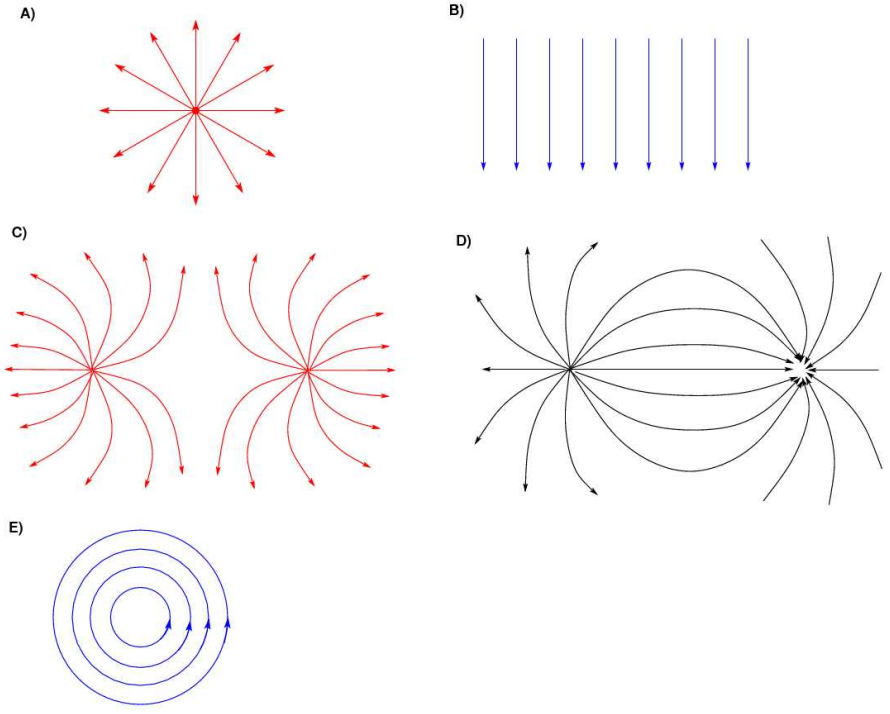
7. Two identical point charges q_1 and q_2 are at a distance r apart. If the size of q_1 is doubled and the distance between them tripled, the strength of the electrical force between them

   (A) goes up by a factor of 3.
   (B) goes down by a factor of 3.
   (C) goes down by a factor of 9.
   (D) goes down by a factor of $2/3$.
   (E) goes down by a factor of $2/9$

8. The below fields represent electric or magnetic fields associated with a long current-carrying wire, a single positive charge, two oppositely charged particles, two positive charges, and the field between two oppositely charged parallel sheets.

   Which list below corresponds to these fields in the order just given?

   (A) A, D, C, B, E
   (B) E, A, D, C, B
   (C) E, D, A, B, C
   (D) D, E, C, B, A
   (E) C, B, A, E, D



9. A positively charged plate sits above a negatively charged plate, as shown below.

An electron is shot off the top plate toward the bottom plate with nonzero initial velocity. The electron's maximum velocity is reached

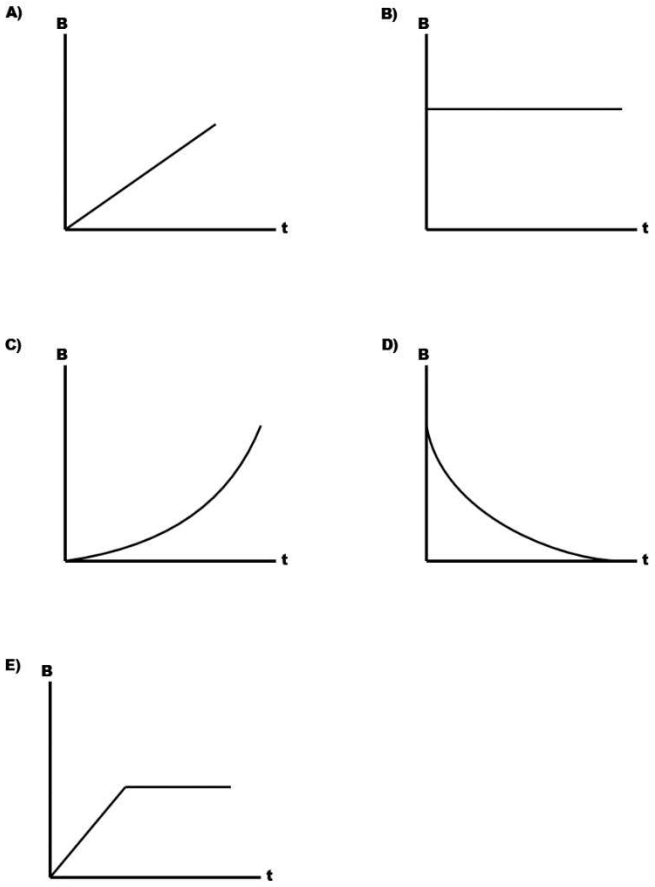
(A) at the bottom plate.
(B) at the top plate.
(C) one-quarter of the distance between the plates.
(D) halfway between the plates.
(E) everywhere; the velocity is constant.

10. An electron is released from rest just above a very large, negatively charged sheet, which carries surface charge density $\sigma = 5\ \text{nC/m}^2$. When the electron is $0.5\ \text{m}$ above the sheet its speed is most nearly

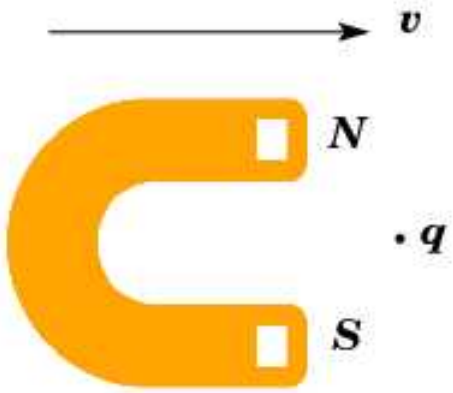
(A) $4 \times 10^8\ \text{m/s}$
(B) $3 \times 10^8\ \text{m/s}$
(C) $5 \times 10^7\ \text{m/s}$
(D) $1 \times 10^7\ \text{m/s}$
(E) $300\ \text{m/s}$

11. Suppose that the current I in a solenoid is directly proportional to time: $I = bt$, where b is a constant. Then the magnetic field inside the solenoid is best described by which graph of B vs t?

(A) A
(B) B
(C) C
(D) D
(E) E



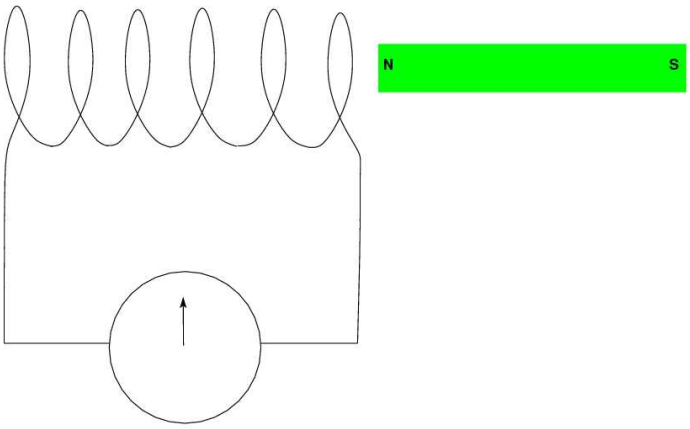
12. A horseshoe magnet with a field of 0.15 T is moving at a constant velocity $\mathbf{v} = 2.5\ \mathrm{m/s}$ to the right and encounters a charge $q = -5$ nC, as shown below.



The direction and magnitude the magnet exerts on the charge is:

(A) 1.875×10^{-9} N, to the right
(B) 1.875×10^{-9} N, into the page
(C) 1.875×10^{-9} N, out of the page
(D) 0.375×10^{-9} N, out of the page
(E) 0.375×10^{-9} N, into the page

13. A permanent bar magnet is at rest such that its north pole is very close to the right end of a coil of wire attached to a meter, as shown below.

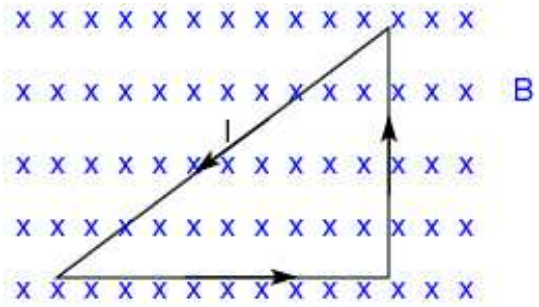


In order to generate an electrical current registered by the meter, you could

i. move the magnet away from the coil at constant velocity
ii. move the magnet into the coil at constant velocity
iii. accelerate the magnet into the coil
iv. increase the area of the coil
v. do nothing—the meter is already registering a current

(A) i
(B) i and ii
(C) i, ii and iii
(D) i, ii, iii and iv
(E) i, ii, iii, iv and v

Questions 14 – 15 refer to the following figure.



14. A current of 7 amps is flowing around the loop shown above. If the magnetic field strength is 0.2 T and the hypotenuse has length 70 cm, the force on it is closest to

(A) 0.14 N
(B) 0.48 N
(C) 0.98 N
(D) 1.44 N
(E) 1.92 N

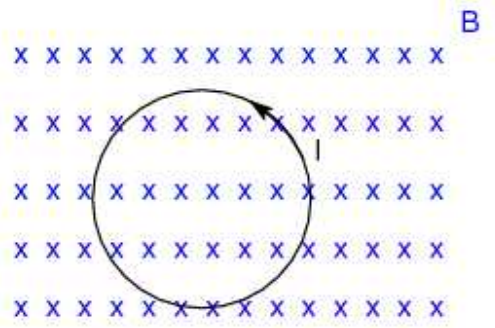
15. The direction of the force on the hypotenuse is

(A) pointed to the north.
(B) to the south.
(C) pointed to the northwest.
(D) pointed to the southeast.
(E) pointed to the southwest.

16. A current I is flowing counterclockwise in a circular loop in a uniform B-field, as shown in the figure.

At 4:00 in the diagram, the magnetic force on the loop is pointed in which direction?

(A) Into the page
(B) Out of the page
(C) Towards 10:00
(D) Towards 12:00
(E) Towards 6:00



Questions 17 – 19 refer to the following figure.

17. Two positive charges of magnitude q and $2q$ are fixed in place along the x-axis. Is there any place along the x-axis where the total field could be zero?

(A) Yes, somewhere to the left of the charge q
(B) Yes, somewhere to the right of the charge $2q$
(C) Yes, between the two charges but closer to q
(D) Yes, between the two charges but closer to $2q$
(E) No, the field can never be zero

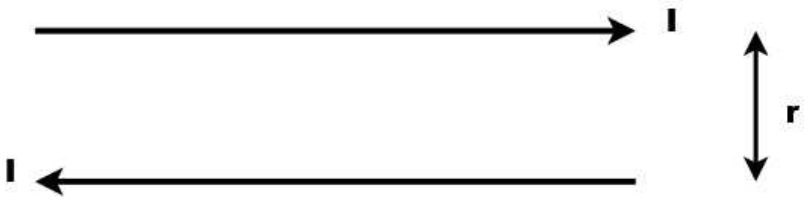
18. The signs of both charges in the previous problem are changed from positive to negative. Is there any point along the x-axis where the electric field could be zero?

(A) Yes, somewhere to the left of the charge marked q
(B) Yes, somewhere to the right of the charge marked $2q$
(C) Yes, between the two charges but closer to q
(D) Yes, between the two charges but closer to $2q$
(E) No, the field can never be zero

19. The sign of q remains positive and the sign of $2q$ is changed to negative. Is there any point along the x-axis where the electric field could be zero?

(A) Yes, somewhere to the left of the charge marked q
(B) Yes, somewhere to the right of the charge marked $2q$
(C) Yes, between the two charges but closer to q
(D) Yes, between the two charges but closer to $2q$
(E) No, the field can never be zero

Questions 20 – 22 refer to the following figure.



20. Two long, parallel wires separated by a distance r carry equal currents I in opposite directions, as shown. The direction of the field caused by the top wire at the position of the bottom wire and the direction of the force exerted by the top wire on the bottom wire are

    (A) B into the page; F down
    (B) B up; F into the page
    (C) B into the page; F up
    (D) B out of the page; F down
    (E) B down; F out of the page

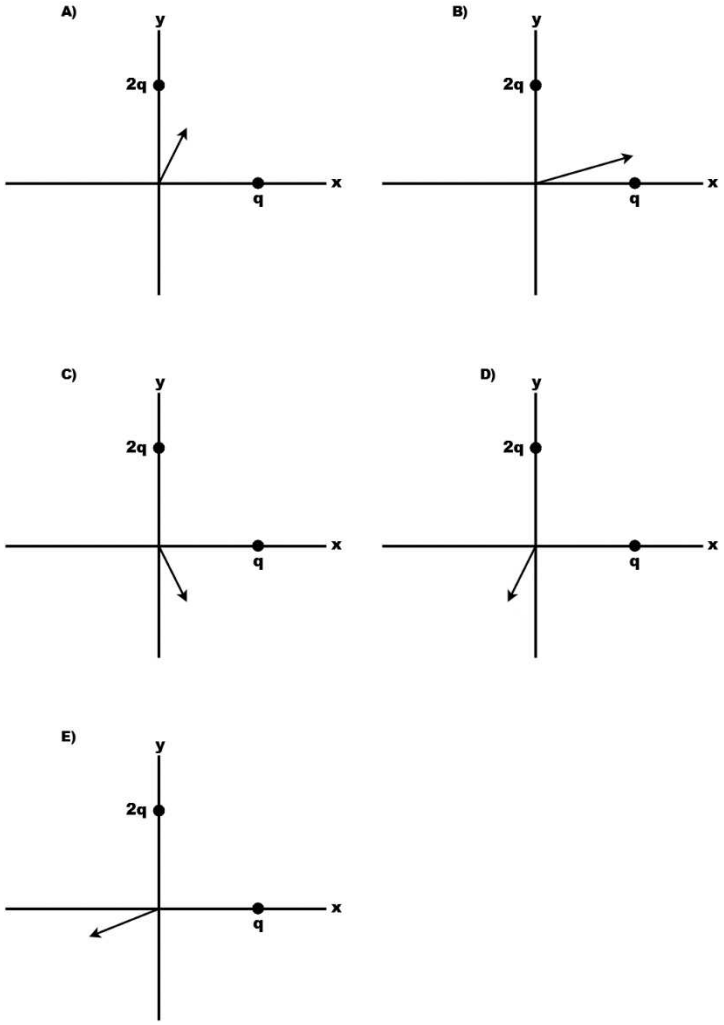
21. In the previous problem, the direction of the magnetic field produced by the bottom wire and the direction of the force exerted by the bottom wire on the top wire are

    (A) B into the page; F down
    (B) B up; F into the page
    (C) B into the page; F up
    (D) B out of the page; F down
    (E) B out of the page; F up

22. The two wires in the previous problem each carry 1 amp of current and are separated by a distance $r = 1$ m. The force per meter between the two wires is

    (A) $4\pi \times 10^{-2}$ N/m
    (B) 2×10^{-4} N/m
    (C) $4\pi \times 10^{-7}$ N/m
    (D) $2\pi \times 10^{-7}$ N/m
    (E) 2×10^{-7} N/m

Questions 23 – 25 refer to the following figure.



23. Two positive charges with magnitude q and $2q$ sit at points $(1,0)$ and $(0,1)$ on the x- and y-axis, respectively. Which figure best represents the total electric field at the origin?

(A) A
(B) B
(C) C
(D) D
(E) E

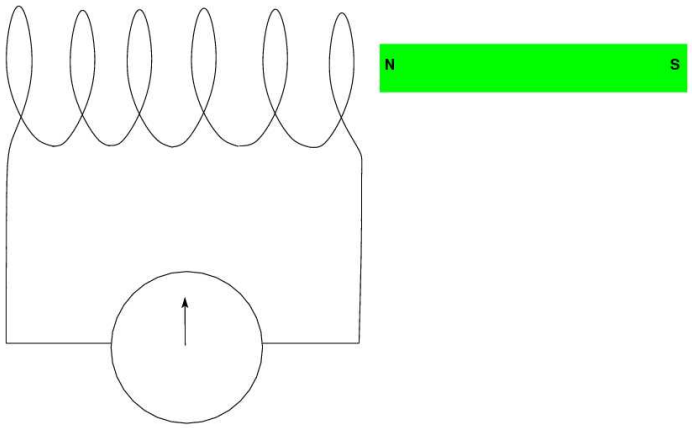
24. Two negative charges with magnitude q and $2q$ sit at points $(1,0)$ and $(0,1)$ on the x- and y-axis, respectively. Which figure best represents the total electric field at the origin?

(A) A
(B) B
(C) C
(D) D
(E) E

25. A negative charge with magnitude q and a positive charge with magnitude $2q$ sit at points $(1,0)$ and $(0,1)$ on the x- and y-axis, respectively. Which figure best represents the total electric field at the origin?

    (A) A
    (B) B
    (C) C
    (D) D
    (E) E

Questions 26 – 27 refer to the following figure.



26. Instead of a meter, a device (such as a battery) is attached to the coil that can generate a current. When the device is turned on the bar magnet will

    (A) remain at rest.
    (B) be attracted to the coil.
    (C) be repelled from the coil.
    (D) shoot through the coil.
    (E) (B) or (C) depending on which way the current is flowing in the coil

27. The magnet is at rest far to the right and then shot into the coil with an initial speed v. As the magnet approaches the coil it will

    (A) keep moving at speed v.
    (B) slow down.
    (C) speed up.
    (D) gain potential energy.
    (E) lose potential energy.

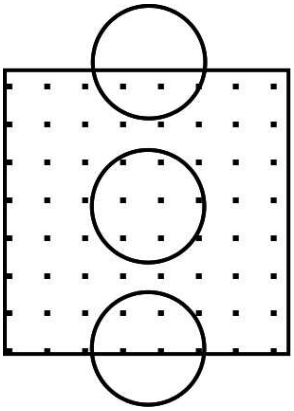
28. A particle of mass m and charge $+q$ is shot with velocity **v** into a region of uniform magnetic field **B**. Suppose that **v** points to the top of the page and **B** points out of the page. Ignoring gravity, the particle

    (A) travels in a straight line.
    (B) moves clockwise in a circle with radius $r = qB/mv$.
    (C) moves counterclockwise in a circle with radius $r = qB/mv$.
    (D) moves clockwise in a circle with radius $r = mv/qB$.
    (E) moves counterclockwise in a circle with radius $r = mv/qB$.

29. Proton A is moving with speed 10^6 m/s in a magnetic field of 0.01 T. Proton B is moving in the same magnetic field with speed 2×10^6 m/s. One can conclude that

(A) Both protons have an orbital radius of about 1 meter.
(B) Both protons have the same orbital period of about 6.5×10^{-6} s.
(C) proton B has twice the orbital period of proton A.
(D) proton B has twice the orbital frequency of proton A.
(E) proton B has half the orbital radius of proton B.

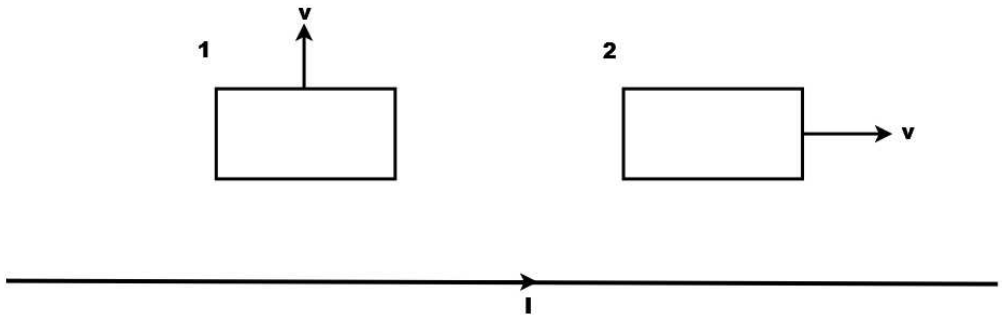
30. A conducting ring is dropped through a region of constant B-field, as shown.



As the ring falls through the B-field, an electrical current is induced in the ring that

(A) flows counterclockwise at the top, is zero in the center and flows clockwise at the bottom.
(B) flows clockwise at every position.
(C) flows counterclockwise at every position.
(D) flows clockwise at the top, is zero in the center and flows counterclockwise at the bottom.
(E) is always zero.

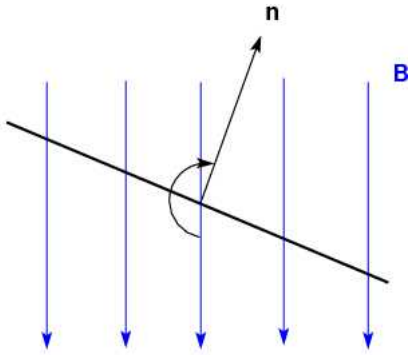
31. Two wire loops, 1 and 2, are moving with equal speeds in the directions indicated below relative to a very long wire carrying a current I.



Which statements are true about the loops? (More than one answer allowed.)

i. A clockwise current is induced in loop 1; ii. A clockwise current is induced in loop 2;
iii. A counterclockwise current is induced in loop 1; iv. A counterclockwise current is induced in loop 2;
v. Zero current is induced in loop 1; vi. Zero current is induced in loop 2

32. A circular loop rotates around a horizontal axis coming out of the page (side view shown, below). The loop is rotating in a uniform B-field, pointed downward.



For every two full rotations, how often does the induced current change direction?

(A) Twice
(B) Four times
(C) Eight times
(D) Twelve time
(E) Sixteen times

Questions 33 – 37 refer to the following figure.

33. Two oppositely charged parallel plates are held apart at distance of 10 cm. The potential difference between them is maintained at 50 volts. The electric field between them is

(A) 5 V/m, down
(B) 50 V/m, up
(C) 500 V/m, down
(D) 500 V/m, up
(E) 1/50 V/m, down

34. A test charge $+q$ and a test charge $-q$ are released midway between the two plates. Let the voltage of the top plate be V and the voltage of the bottom plate be 0. The distance between the two plates is d. Which of the following statements about the test charges are true?

    i. Both charges have gained a kinetic energy of $|q|Ed/2$ when they hit the plates.
    ii. Charge $+q$ is initially at a positive voltage and charge $-q$ is initially at a negative voltage .
    iii. Charge $+q$ initially has a positive potential energy and charge $-q$ initially has a negative potential energy.
    iv. Both charges lose potential energy.

    (A) i only
    (B) i and ii
    (C) i and iii
    (D) iii and iv
    (E) i, iii and iv

35. At what initial velocity v_o must an electron be shot off from the top plate in order to reach the bottom plate?

    (A) 2.1×10^3 m/s
    (B) 2.1×10^5 m/s
    (C) 4.2×10^6 m/s
    (D) 0 m/s
    (E) 4.2×10^8 m/s

36. What potential difference between the two plates would be needed to accelerate a hydrogen ion (a proton) from rest to a speed of 10^6 m/s?

    (A) 100 V
    (B) 500 V
    (C) 1000 V
    (D) 5000 V
    (E) 10,000 V

37. The voltage in the previous problems is maintained at 50 V. An uncharged copper block of thickness 2 cm and length equal to that of the plates is inserted exactly midway between them. What is the field strength between the bottom of the copper block and the bottom plate?

    (A) 0 V/m
    (B) 250 V/m
    (C) 500 V/m
    (D) 625 V/m
    (E) 2000 V/m

38. Metal sphere A has a radius of 5 cm and metal sphere B has a radius of 10 cm. Sphere A carries a charge of 9 nC and sphere B carries a charge of 18 nC. If the surfaces of A and B are 185 cm apart, the potential energy between them is

    (A) 7.29×10^{-17} J
    (B) 7.88×10^{-9} J
    (C) 7.88×10^{-7} J
    (D) 7.29×10^{-9} J
    (E) None of the above

39. A third sphere carrying a charge of 9 nC is brought to 2 m of the center of sphere A in the previous problem and to 3 m of the center of sphere B. The total potential energy of the system is now

    (A) 1.58×10^{-9} J
    (B) 7.29×10^{-7} J
    (C) 1.21×10^{-6} J
    (D) 1.58×10^{-6} J
    (E) 1.58×10^{-5} J

Free-response problems

40. Two masses of 10^{-9} kg each carry a charge of $+3$ nC and are initially held motionless by a massless thread at a distance of 9 m.

    a) Draw the electric field produced by the two charges.

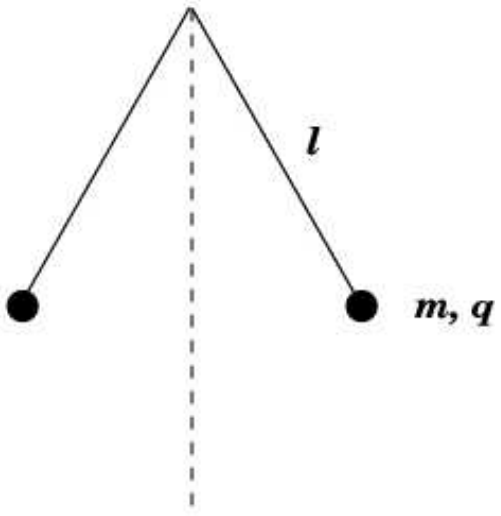
    b) What is the tension in the thread?

    c) The thread is cut by a magic wand. What is the initial acceleration of each of the charges?

    d) How does the acceleration change with r? Graph the behavior. Explain qualitatively how the velocity will change with r. In particular, what happens to the velocity when r goes to infinity. Where is the maximum velocity?

41. Two identical plastic balls of mass 10 gm each are hung by threads with length 30 cm from a common point, as shown below.

The balls are each charged with the same charge q and repel each other until they come to rest with a horizontal separation of 30 cm.



a) Sketch the electric field produced by the two balls.

b) Draw the force vectors on the right-hand ball.

c) What is the charge q in each ball?

42. A proton is launched from a very long negatively charged plate at an initial velocity v_o and angle θ toward an identical, positively charged plate, as shown below.

The two plates are held at a potential difference $V = 1000$ volts, and the distance between the two plates is d.

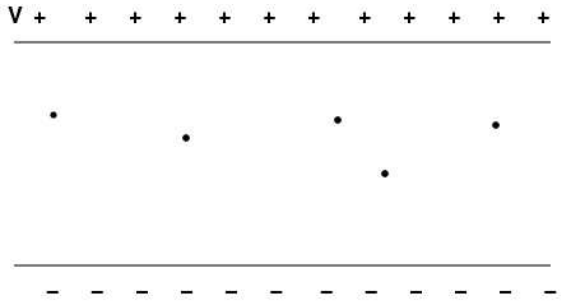
a) Draw in the electric field vectors. Write an algebraic expression for the electric field strength in terms of given quantities.

b) If $v_o = 4 \times 10^5$ m/s, and $\theta = 60°$, does the proton hit the top plate? If so, with what velocity?

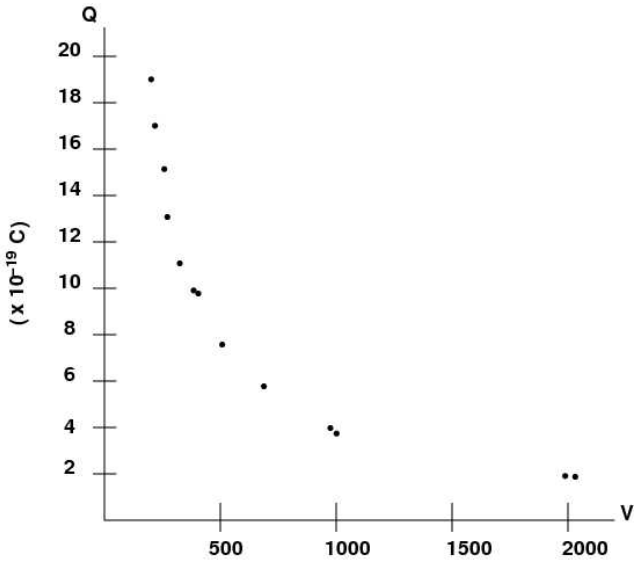
c) If $d = 25$ cm, how long is the proton in flight?

d) If a uniform magnetic field pointing along the proton's initial velocity vector is introduced between the plates, does this alter the conclusion to (b)?

43. The Millikan oil-drop experiment was the first experiment that attempted to determine the charge on the electron. Today students perform modern versions of the experiment in various ways, usually involving small latex spheres. The spheres are injected between two conducting plates held at a potential difference V. Assume that the plates themselves are contained in a vacuum chamber.



The spheres are of identical mass but may carry different and unknown amounts of charge. The voltage is adjusted until a selected sphere rises a constant velocity and the calculated charge on the sphere is recorded. This procedure is repeated many times for many different spheres. Eventually a graph of the results is produced:



a) Write down the condition that the spheres rise with constant velocity as a function of voltage and charge.

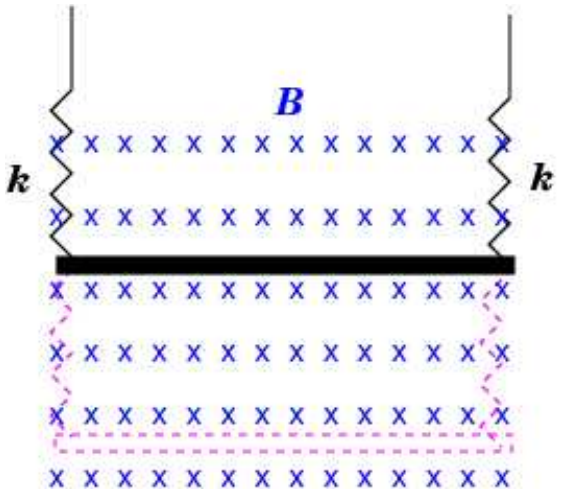
b) What shape should the theoretical curve of the charge versus the voltage be?

c) How might you physically account for the lack of data between 1000 V and 2000 V?

d) If the distance between the plates is 5 cm, what is the mass of the latex spheres?

44. A conducting wire of length 0.5 m and mass 10 gm is suspended in a uniform magnetic field of 0.5 T by two identical conducting springs with spring constant k (see below).

A current of 1 amp is sent through the springs and wire, and the wire is observed to move downward 1 cm, stretching the springs.



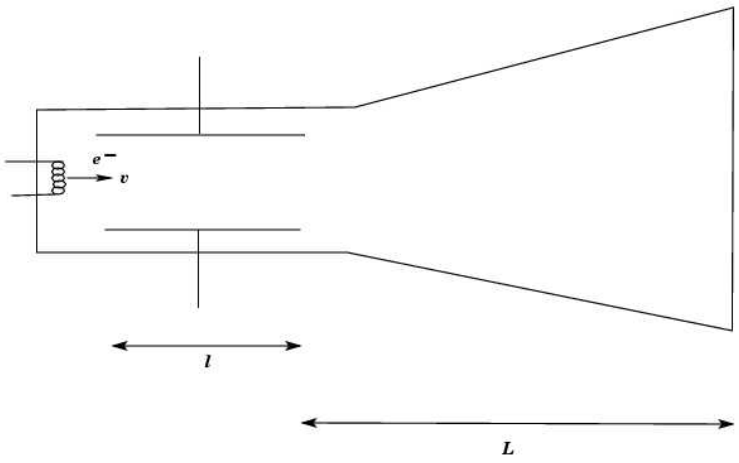
a) In which direction is the current flowing?

b) Derive an expression for the spring constant in terms of known quantities and fundamental constants.

c) What is the numerical value of k?

45. Until recently, TV sets consisted of what was called a cathode-ray tube. Electrons were boiled off a light bulb filament and passed through plates that deflected the particles; the electron beam went on to strike the TV screen itself and lit up a pixel of phosphor. The beam could be deflected both up and down and sideways (only one direction shown). By deflecting the electron beam and scanning it across the screen a picture was built up. Consider the simplified TV below.

An electron beam is shot along the centerline between the two deflector plates. The top plate is held by a power supply (not shown) at a voltage $V = +1000$ volts above the bottom plate. The separation between the plates is $d = 4$ cm. The length of the plates is $\ell = 8$ cm. The distance L from the end of the deflector plate to the TV screen is 30 cm. The electrons enter the region between the deflector plates with a velocity $v = 4 \times 10^7$ m/s.



a) In which direction is the electric field between the deflector plates?

b) In which direction are the electrons deflected?

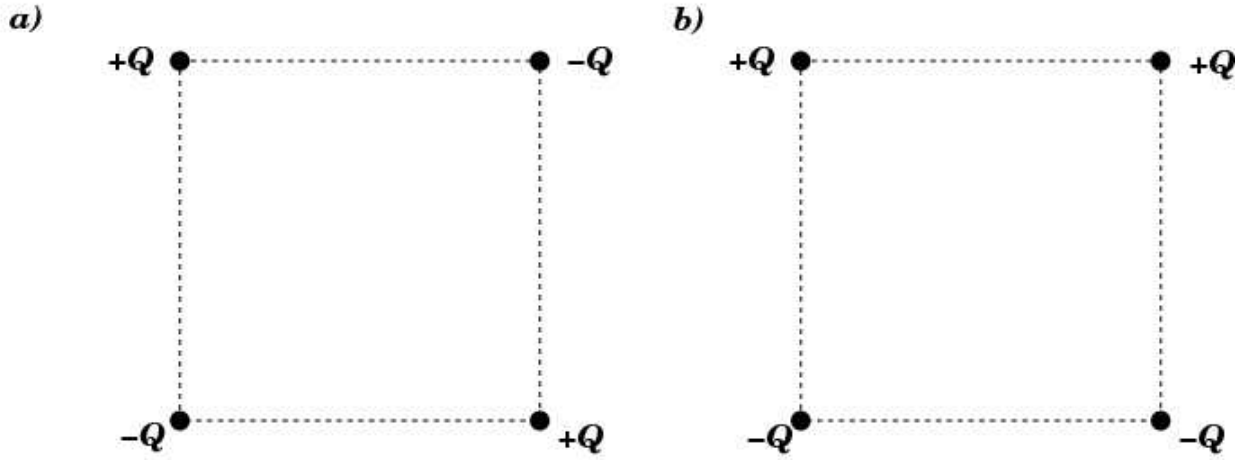
c) What is the force on an electron? (Neglect gravity.)

d) What is the acceleration on the electron?

e) What is the total deflection of the electron as it hits the TV screen?

f) What size magnetic field and in what direction would you need to place between the deflector plates in order to prevent the electrons from being deflected?

46. Four charges of equal magnitude are arranged in a square in two configurations, as shown below. The side length of the square is s.



a) What is the magnitude and direction of the electric field at the center of the square in configuration (a)?

b) What is the magnitude and direction of the electric field at the center of the square in configuration (b)?

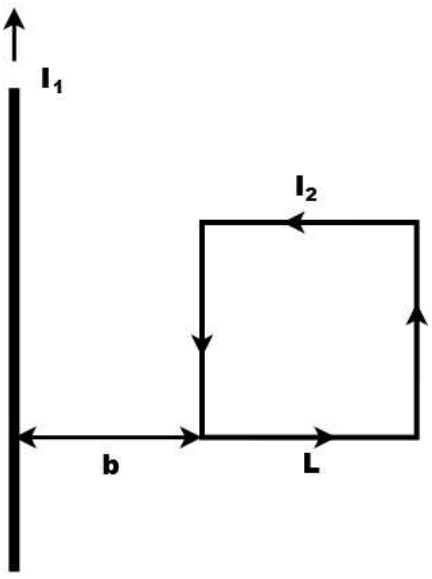
c) If the top left charge in (a) is removed, what is the electric field (magnitude and direction) at that point due to the other three charges?

d) What is the total potential energy of configuration (a) with all four charges present?

e) What is the total potential energy of configuration (b) with all four charges present?

f) For which configuration would it take more work to remove the charge from the lower right corner to an infinite distance from the square?

47. A vertical wire carries a current I_1 up in the plane of the page, as shown below. A square wire loop with side length L sits in the plane of the page at a distance b from the wire and carries a current I_2 in the counterclockwise direction.



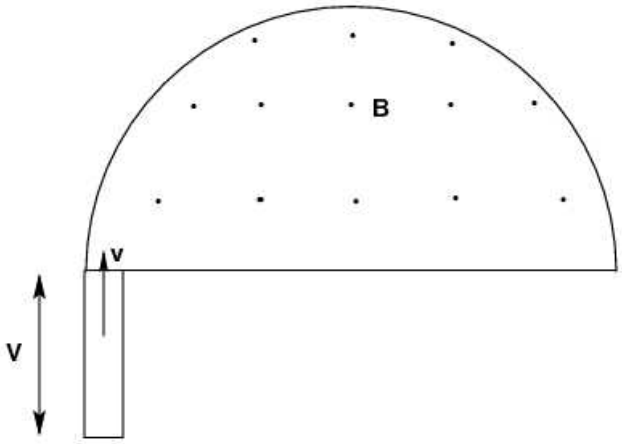
a) Is there a net force on the loop? If so, in what direction?

b) Is there a net torque on the loop around a vertical axis running through the loop's center? If so, what is its action on the loop?

c) If $I_1 = 2$ A; $I_2 = 1$ A; $b = 10$ cm; $L = 5$ cm, what is the magnitude of the net force?

d) What is the magnitude of the net torque?

48. A mass spectrometer is a device for measuring the mass of atomic isotopes. As shown below, it consists of a small accelerator, which accelerates singly-ionized atoms across a voltage V. (Singly-ionized atoms have had one electron removed.) The ions are injected into a region with a constant magnetic field **B** (here pointing out of the page).



a) Assuming the ions are accelerated from rest, with what velocity v are they injected into the magnetic field? (Write v in terms of the voltage V, the charge on the ion q and the ion's mass m.)

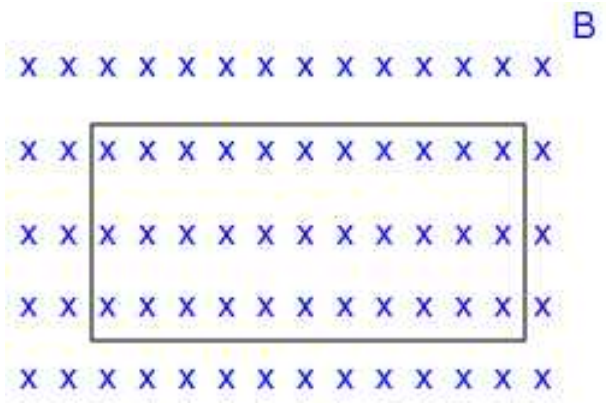
b) What is the magnitude and direction of the force **F** acting on the ions?

c) On the diagram, draw the trajectory that the ions follow.

d) Write an algebraic expression involving the above that gives the distance d from their point of entry where the ions end up.

e) You are given a sample consisting of ^{24}Mg and ^{26}Mg (magnesium). The former atoms have mass $m = 3.983 \times 10^{-26}$ kg. If $V = 2500$ V, $B = 556$ G, how far apart will the atoms end up?

49. A rectangular loop is situated in a region with a uniform magnetic field of 0.1 T pointing into the page, as shown below. The length of the loop is 20 cm and the width is 10 cm.



a) What is the magnetic flux through the loop?

b) If the value of B is increased from 0.1 T to 0.5 T in 0.3 s, what will be the EMF induced into the loop? What will be the direction of the induced current? Justify your answer.

c) If the resistance of the loop is $R = 1\ \Omega$, what is the value of the current?

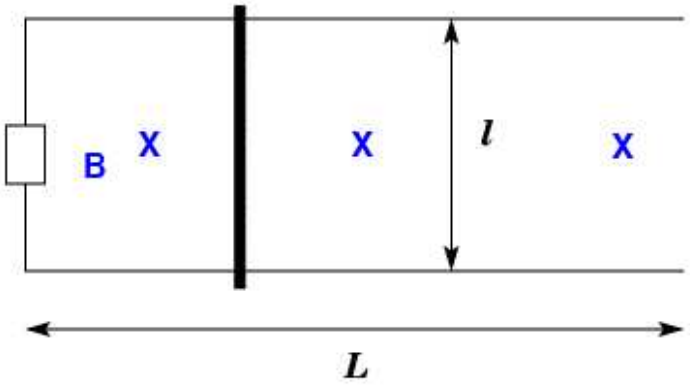
d) If the B-field is tilted at an angle $\theta = 15°$ from the normal to the page, what is the flux through the loop?

e) Give two ways that one could change the induced EMF in the loop, other than changing B.

f) Draw a rotation axis from left to right in the plane of the page that passes through the loop's center. The loop is rotated about this axis at a constant angular velocity. Write an expression for the flux in terms of the angular velocity ω and the time t.

50. The rail gun, currently being developed by the US Navy, is a weapon designed to shoot a projectile off a set of rails by electromagnetic means. Consider a simplified version of the gun, below.

A power supply (box) sends a current I around a circuit consisting of two frictionless rails (horizontal lines) and through a projectile (vertical bar) of length ℓ and mass m. The total length of the rails is L. A uniform magnetic field, B, covering the entire area of the system, is directed into the page.



a) In what direction must the current I be directed to make the projectile move to the right?

b) What is the magnitude of the force on the projectile that directs it to the right?

c) Are there any other forces on the projectile? If so, what is their direction and magnitude?

d) Considering only the force in (b), what is the velocity of the projectile as a function of position from the left end of the rails?

e) If $L = 10$ m; $\ell = 0.1$ m; $m = 2$ kg; $I = 10^4$ A, what is the launch speed of the projectile?

51. A wire carries a current of $I = 10$ A into the page. With a magnetic field probe, a student measures the B-field at five points and notes down the results: $(1, 2.2)$; $(2, 0.85)$; $(5, 0.35)$; $(8, 0.31)$; $(10, 0.11)$. The first number in each pair is the distance to the right of the wire in cm and the second number is the B-field in units of 10^{-4} T.

a) Plot the data, then plot the expected curve on the same graph.

b) What would you say is the average error of the measurements?

c) After you have plotted the data, someone reports that the earth's B-field at this location is 0.5 gauss, pointing to the top of the page. Replot the corrected B-field.

d) What percentage error does the earth's field introduce to the measurements as a function of r?

e) Would you say the original results were trustworthy?

52. A solid copper sphere of radius $R = 10$ cm carries a charge of $+5$ nC.

a) Give an algebraic expression for the electric field at all distances $r < R$. What is the numerical value of the field at $r = 5$ cm?

b) Give an algebraic expression for the electric potential at all distances $r < R$. What is the numerical value of the potential at $r = 5$ cm?

c) Sketch a graph of both the field and potential for $r \geq 0$.

d) A second, uncharged copper sphere of radius $r = 5$ cm is brought near the first sphere. Graph the electric potential everywhere from $0 \leq r \leq \infty$.

10. SIMPLE CIRCUITS REVIEW

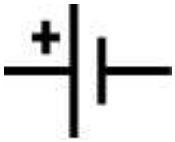
The material covered by this chapter appears on both the SAT subject test and the AP exams. Sections 10.1-10.4 are at the subject-test level. Sections 10.5-10.8 are at the AP level.

1. Essential Terms

The concepts of **charge**, **electric field**, **current** and **voltage** from Chapter 9 become the basis for simple **direct current (DC) circuits**, which are an essential part of daily life and SAT exams. As mentioned in Section 9.7, a **circuit** is a collection of wires and electrical components connected in such a way that an electric current can run through all of them and return to its starting point. Most elementary circuits rely on three essential components: **batteries**, which supply a voltage and drive the current, **resistors**, which impede electrical current, and **capacitors**, which store charge and electrical energy. An equally essential circuit element is the **inductor**, a small solenoid that stores magnetic energy, but since these are not covered on the College Board exams we shall also neglect them. The entire power industry relies on **alternating current** (AC) in which the voltage and current oscillate sinusoidally, but AC power is also not part of the College Board exams and so we spurn it as well.

2. Batteries

Any circuit requires a supply of electric current, which is generally termed the **power supply**. In simple circuits the current I is usually supplied by a **battery**, a device that changes chemical energy into electrical energy. The electrical symbol for a battery is:



The positive end of the battery (+) is the end from which the positive, conventional, current comes out, although as discussed in Section 9.7, it is really electrons that are moving; they emerge from the negative end. Any battery produces a voltage V. This is really the voltage (electric potential) difference between the positive and negative ends of the battery. Since voltage is a potential *difference*, it is always measured *across* a component—from one end to the other. Scientists and engineers usually refer to the voltage difference across the component as a **voltage drop**. A battery rated at 12 V has a 12-volt potential difference between the positive and negative ends, and by energy conservation any circuit connected between the two ends must also drop 12 volts.

Any circuit must be **closed**: the current must be able to complete a complete round trip (hence the term *circuit*.) If it is not closed, the current simply cannot flow.

Example 10.1: Short circuits. The simplest circuit is a battery whose positive end is connected to its negative end through a switch, S:

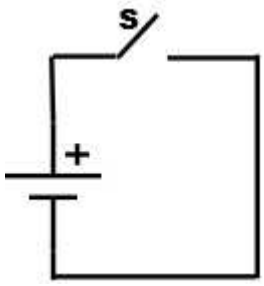


Figure 10.1

This is a circuit you want to avoid constructing, because when S is closed you have created a **short circuit**: a large current immediately flows between the positive and negative end of the battery, which can lead to overheating and the destruction of the battery.

3. Resistors

A **resistor** is a device that impedes the flow of electrical current. It is often useful to regard electrical current as analogous to water flow and resistors as analogous to pipes. A wide pipe allows a large flow rate and has a small resistance, whereas a narrow pipe has a large resistance and allows only a small flow rate. Electrical resistors are typically small cylinders made of carbon or other material that conduct electrical current poorly. The electrical symbol for a resistor is

Resistors are necessary, among other things, to avoid short circuits of the type in Example 10.1. When a resistor is inserted into the above circuit it is found that the voltage drop across the resistor is directly proportional to the current flowing through it, a famous result known as

Ohm's Law

$$V = IR \quad (10.1)$$

The constant of proportionality between the voltage and current, R, is termed the **resistance** of the component and is a measure of how much it impedes electrical current: Since $I = V/R$, the higher R for a given voltage, the lower the current. The SI unit of resistance is the ohm (Ω), after Georg Simon Ohm. $1\ \Omega \equiv 1\ \mathrm{V}/1\ \mathrm{A}$.

Ohm's law is not a law of nature. It is an experimental result that may be regarded as the definition of a resistor: If a material or electrical component obeys Ohm's law, it is a resistor. If it does not, it is not a resistor.

Ohm's law can be illustrated by Figure 10.2:

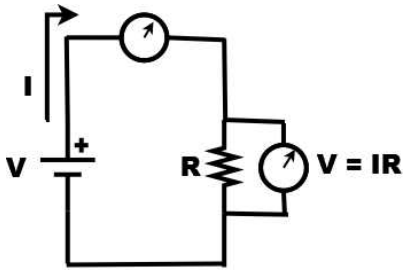


Figure 10.2

The battery sends a current I, which is measured by a current meter (an **ammeter**), *through* the resistor R. At the same time, you measure the voltage *across* the resistor by a **voltmeter**. If the battery voltage can be adjusted, the current I will also change. Plotting V as a function of I, you find simply $V = IR$, a straight line with slope R, as in Figure 10.3.

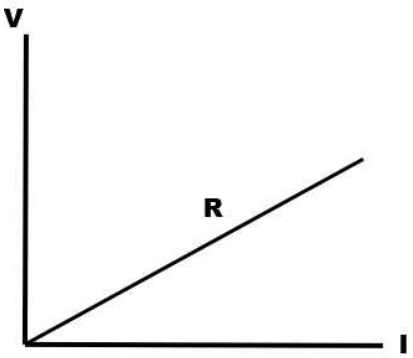


Figure 10.3

If the voltmeter is removed from Figure 10.2 and the ammeter is replaced by a wire, we have merely a battery connected to itself through a resistor. The positive end of the battery is at voltage V and the negative end is at 0 volts. The charges at the positive end therefore have potential energy $U = qV$ and at the negative end have $U = 0$. The top end of the resistor also has voltage V and the bottom end is at zero volts. Because the full current is flowing through the resistor, the entire energy of the charges must have disappeared there. However, by conservation of energy, the energy cannot have disappeared. Where has it gone? *It is dissipated by the resistor as heat*.

From Section 3.10 we know that power is the expenditure of energy per unit time and that the SI unit of power is the watt (W). $1\ \text{W} \equiv 1\ \text{J/1s}$. Note that the dimensions of power are energy/time. By definition, an electrical current I has dimensions charge/time, while by definition voltage has dimensions of energy/charge. Thus, $IV = \frac{\text{charge}}{\text{time}} \times \frac{\text{energy}}{\text{charge}} = \frac{\text{energy}}{\text{time}} = \text{power}$. Hence, the power transmitted by an electrical current I though a voltage V is

$$P = IV \quad (10.2)$$

Eq. (10.2) is the definition of electrical power and is true in any circuit, whether a resistance is present or not. For example, if an electron beam in a TV (Problem 9.45) has a current of 1 A and is shot across a potential difference of 2000 V, then the power of the beam is 2000 W, or 2 kW. *In a resistor* Ohm's law applies and $V = IR$. Substituting for V or I in Eq. (10.2) shows that

$$P = I^2R \quad (10.3a)$$

$$P = \frac{V^2}{R} \quad (10.3b)$$

The basis of modern civilization

This is the power dissipated as heat in any resistor; it is usually termed **Joule heating** or $\boldsymbol{I^2R}$ **losses**. In other words, it is the energy lost in power transmission lines or any electrical appliance in which a resistance is present.

Example 10.2: Power losses. The average power consumption for a house in Anytown USA is about 1200 W, or 1.2 kW. If Anytown has 1000 homes, the total power consumption is 1.2×10^6 W, or 1.2 MW. Power is usually transmitted at high voltage, above 10^5 V. Why such high voltage?

From Eq. (10.2), the power Anytown requires is $P = IV = 10^5 I$. If $P = 1.2 \times 10^6$ W, then $I = 12$ A. Assume there are ten power lines going into Anytown with each line carrying 1.2 A. Power lines have resistance. If a transmission line to Anytown has a resistance of 10 Ω, then Eq. (10.3a) shows that for $I = 1.2$ A, the transmission losses for each line are 14.4 W, for a total of 144 W. This is only 0.012% of Anytown's power requirements.

If instead power were transmitted at 1000 V, to supply Anytown with the same 1.2 MW would require a current 100 times larger: $I = 1200$ A, or 120 A per line. The power losses are proportional to I^2; if the current goes up a factor of 100 then the power losses per line go up by a factor of 10^4 to 1.4×10^5 W, for a total power loss of 1.4×10^6 W—more than 100% of Anytown's power requirements. This is why high voltage is used to transmit power.

4. Resistors in Series and Parallel

One regularly needs to connect resistors together in a circuit and it becomes essential to know how the individual resistances combine. Any two components of a circuit are said to be in **series** when they are attached "end to end" so that the same current passes through both of them:

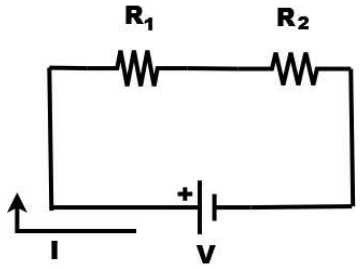


Figure 10.4

In a series circuit the current is the same everywhere.

We can find the total resistance of this circuit as follows. If there is a voltage V across the battery, then energy conservation demands that the voltage drop across the two resistors equals the battery voltage: $V = V_1 + V_2$, where V_1 and V_2 are the voltage drops across R_1 and R_2, respectively. Since the current I is the same through both resistors, Ohm's law applied to each gives $V_1 = IR_1$ and $V_2 = IR_2$. Then $V = IR_1 + IR_2 = I(R_1 + R_2)$. This can be rewritten as $V = IR_{eq}$, where $R_{eq} = R_1 + R_2$ is the total, or **equivalent resistance**, of the circuit. Thus,

Resistances in series add:

$$R_{eq} = R_1 + R_2 + R_3 + \cdots + \quad (10.4)$$

One can think of the equivalent resistance as the resistance of a single resistor that can replace all the others in a circuit without affecting the total current.

Two electrical components are said to be in **parallel** when they sit "side by side" in a circuit and *the voltage across them is the same*:

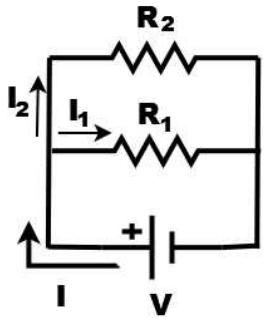


Figure 10.5a

In a parallel circuit the current divides at a junction. In Figure 10.5a, I is the total current coming out of the battery. The part of I labeled I_1 goes into R_1 and the part labeled I_2 goes into R_2. Because by conservation of charge current cannot disappear—just as water in a pipe can't disappear—we must have $I = I_1 + I_2$.

The equivalent resistance can now be found. Each resistor above is connected across the *entire* battery voltage V. Ohm's law then requires $V = I_1 R_1$ and also $V = I_2 R_2$. Then $I_1 = V/R_1$ and $I_2 = V/R_2$. As in the series circuit, we imagine replacing R_1 and R_2 by a single resistor with resistance R_{eq}:

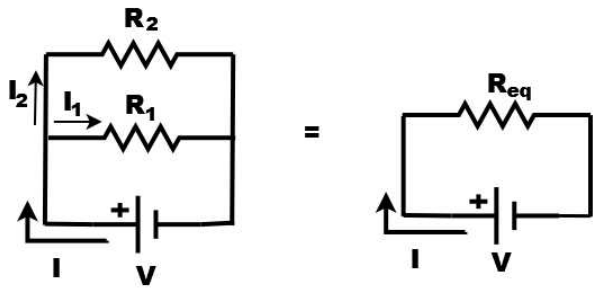


Figure 10.5b

Since I is the total current going into R_{eq}, Ohm's law gives $I = V/R_{eq}$. The condition $I = I_1 + I_2$ becomes from above

$$\frac{V}{R_{eq}} = \frac{V}{R_1} + \frac{V}{R_2}.$$

Cancelling the V's shows that

In parallel circuits inverse resistances add:

$$\frac{1}{R_{eq}} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} + \cdots + \qquad (10.5)$$

Warning: After adding the inverse resistances, remember that you've got to take the inverse once more to get the equivalent resistance, R_{eq}.

Example 10.3: Typical resistor problem. If in Figure 10.6a, $R_1 = 1200\ \Omega$, $R_2 = 500\ \Omega$ and $R_3 = 700\ \Omega$, what is the equivalent resistance in the circuit?

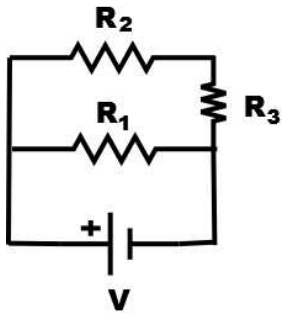


Figure 10.6a

We first observe that R_2 and R_3 are in series because any current going through R_2 also goes through R_3. The equivalent resistance for these two resistors is $R_{23} = R_2 + R_3 = 1200\ \Omega$. In the above diagram we then replace R_2 and R_3 with a single resistor R_{23}:

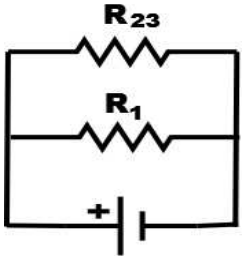


Figure 10.6b

We see that R_{23} is in parallel with R_1. Thus, the equivalent resistance for the entire circuit is given by $1/R_{eq} = 1/R_1 + 1/R_{23}$, or $1/R_{eq} = 1/1200 + 1/1200 = 2/1200$. Therefore $R_{eq} = 600\ \Omega$.

Notice the answer is half the resistance of the individual resistors in parallel.

Tip: The equivalent resistance of two identical resistors in parallel, each with resistance R, is $R/2$.

Expanded tip: For a series of resistors in parallel, $R_1, R_2, R_3 \ldots$, the equivalent resistance must be less than the value of the smallest resistor. (**Exercise:** Show this.) This tip also provides a good reality check for you answers; if you have two resistors $R = 5000\ \Omega$ and $R = 7000\ \Omega$ in parallel, you know before doing any work that the equivalent resistance must be less than $5000\ \Omega$.

5. Light Bulbs

Many SAT problems substitute light bulbs for resistors. Technically, a light bulb does not behave exactly like a resistor because the resistance of the filament changes as the bulb heats up, but if the temperature is constant, then Ohm's law applies and the bulb can be thought of as a resistor.

Usually such problems ask whether the brightness of a certain bulb will change when some aspect of the circuit is altered. For old-fashioned incandescent bulbs, **the brightness of a bulb is proportional to the power going through the filament: $P = I^2R$**. Thus, if the current through the bulb changes, the brightness also changes.

Example 10.4: Typical light bulb problem. The circuit in Figure 10.7 consists of three identical light bulbs, a battery with voltage V, and a switch, S, connected as indicated. (The internal wiring of the bulbs is not shown.) S is initially in the closed position. After it is opened, the brightness of which bulbs will change?

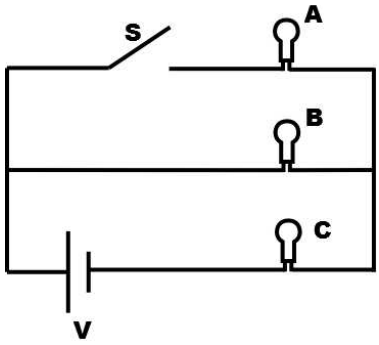


Figure 10.7

This problem actually appeared on an SAT subject test and is a bit tricky; the College Board's explanation of the answer was no more than a "plausibility argument" and not rigorous. We will try to be more convincing.

Because the brightness of a bulb depends on the current passing through it, the question becomes, "Which currents change when the switch is thrown?" When S is opened, bulb A is taken out of the circuit altogether and goes dark. Clearly, the brightness of A changes. What about B and C? This is less clear. With S closed the circuit has an equivalent resistance R_{eq} that depends on all the bulbs. The total current coming out of the battery is $I = V/R_{eq}$, but only part of I goes through each bulb. Because A is initially in parallel with B, the equivalent resistance is *lower* with S closed than with S open. Since R_{eq} goes up when S is opened, the battery current $I = V/R_{eq}$ goes down. With S open, we have a series circuit, so *all* of the reduced current now flows through B and C. However, this smaller current might be the same as the fraction of the larger battery current that flowed through B and C when S was closed. The College Board simply does not explain whether this is so.

Let us be rigorous. Call the resistance of each bulb R. With S closed, bulbs A and B are in parallel; their equivalent resistance is $R_{AB} = R/2$. The resistance R_{AB} is in series with bulb C and so the equivalent resistance of the entire circuit is $R_{eq} = R + \frac{R}{2} = 3R/2$. By Ohm, the total current coming out of the battery is $I_{closed} = V/R_{eq} = 2V/3R$.

Since bulbs A and B are identical, I_{closed} must be split evenly between them, and so the current through B is $I_{B_{closed}} = V/3R$.

With S open, $R_{eq} = 2R$ and the current going through bulb B (and C) is $I_{B_{open}} = V/2R$. As predicted, R_{eq} has gone up and I has gone down. Because $I_{B_{open}} \neq I_{B_{closed}}$, the brightness of bulbs B and C must also change. Thus, all three bulbs do change brightness.

If you think that it is obvious that the brightness of the bulbs changes, see problems 18-20.

6. Resistivity

The resistance of a resistor depends not only on the material from which it is made but on its length and area. Intuitively, you expect that a longer resistor should have a higher resistance but a wider resistor should have a lower resistance, because more current can flow through it. This turns out to be true. One can show that under certain assumptions

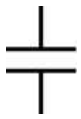
$$R = \frac{\rho L}{A} \qquad (10.6)$$

where L is the length of the material through which the current flows, A is the cross-sectional area through which the current flows and ρ is a property of the material itself, called the **resistivity**. From Eq. (10.6) the dimensions of resistivity are ohms × length and the SI unit is Ω m. Most common metals have ρ in the range of 10^{-8} Ω m, while insulators have ρ up to twenty orders of magnitude higher.

Example 10.5: Power line resistance. What is the resistance of an aluminum power line 100 km long with a radius of one centimeter? The resistivity of aluminum is 2.65×10^{-8} Ω m. For the power line we have $L = 10^5$ m and $A = \pi(0.01)^2$. Plugging these numbers into Eq. (10.6) gives $R = 8.4$ Ω, which is why we chose 10 Ω for the power line into Anytown in Example 10.2

7. Capacitors

A **capacitor** is a device that stores charge and electrical energy. The electrical symbol for a capacitor is



which not coincidentally resembles the two charged plates of Figure 9.4. **Capacitance** is the measure of the amount of charge a capacitor can hold at a given voltage. The larger the capacitance (from the word "capacity") the more charge the device can store.

To motivate the definition of capacitance, consider the simplest capacitor: a charged sphere. From Eq. (9.14), if a test charge q sits at the surface of a sphere with radius R, then the potential energy between q and the sphere is $U = kqQ/R$, where Q is the sphere's charge. This is the amount of work it takes to bring q from a distance of infinity to R. The voltage at the surface of the sphere with respect to infinity is by Eq. (9.15) $V = kQ/R$. Equivalently $Q = VR/k$, implying that for a fixed V, a larger R means a larger Q. This makes intuitive sense; we expect that a larger sphere should be able to hold a larger amount of charge. Let us tentatively define capacitance as "stored charge per volt":

$$C \equiv \frac{Q}{V} \quad (10.7)$$

The reason for defining capacitance in this way is that a) it accords with our expectation that the greater the amount of stored charge, the larger the capacitance; and b) like resistance, it depends only on the device itself, not on Q or V. For the sphere, $C = R/k$. Thus, C is a function entirely of the sphere's geometry.

Note: All voltages are voltage differences. Thus, one should keep in mind that in actuality $C = Q/\Delta V$, where ΔV is the voltage between two points.

Usually when discussing capacitors, texts employ the notation $\epsilon_o \equiv 1/4\pi k$ from Section 9.5, where $\epsilon_o = 8.85 \times 10^{-12}$ in MKS units. Thus, the capacitance of the sphere becomes $C = 4\pi\epsilon_o R$. **All capacitances have dimensions $\epsilon_o \times$ length.**

Consider the positively and negatively charged plates of Section 9.5. From Eq. (9.5) the electric field above or below each plate is $E = 2\pi k\sigma$, where σ is the charge per unit area. Hence, $E = 2\pi kQ/A$, which can be written as $E = Q/2\epsilon_o A$. If we put the positive plate atop the negative plate the field between them is doubled, giving $E = Q/\epsilon_o A$. The voltage between the plates is by definition $V = Ed$ and so $V = Qd/\epsilon_o A$. The capacitance of two charged plates is then from Eq. (10.7),

$$C = \frac{\epsilon_o A}{d} \quad (10.8)$$

The configuration of a positively charged plate near a negatively charged plate is termed a **parallel-plate capacitor.** It is the primordial capacitor. Note that the dimensions of the parallel-plate capacitor are, as for the sphere, $\epsilon_o \times$ length. And again C depends entirely on the geometry of the device: the larger the area of the plates, the greater the capacitance; the smaller the separation, the greater the capacitance.

Tip: All capacitors, regardless of their shape, have a capacitance approximately that of a parallel-plate capacitor.

The SI unit of capacitance is the farad, after Michael Faraday. 1 F $\equiv$ 1 C/1 V.

Example 10.6: Estimate the capacitance of your state without looking anything up. Assume that your state is a parallel-capacitor with plates separated by a distance $d = 1$ m.

Take New Jersey. Residents of NJ know that about three hours are required to drive from the northern tip of the state to Cape May, the southern tip. At 100 km/h, this suggests that the length of the state is approximately 300 km or 3×10^5 m. To drive to the shore from Philadelphia, however, takes only about an hour, suggesting that the width of the state is roughly 100 km, or 10^5 m. We thus estimate the area of NJ to be 3×10^{10} m^2 . From Eq. (10.8), a parallel-plate capacitor of this area and $d = 1$ m has capacitance

$$C \approx 9 \times 10^{-12} \times 3 \times 10^{10} = 0.27 \text{ F}.$$

That NJ's capacitance is less than 1 F demonstrates that a farad is an enormous unit. Until recently, it was very rare to encounter a 1 F capacitor, and if you did, you probably would not live to tell the tale. With nanotechnology, d can be made extremely small and 1 F capacitors about the size of a thumb are now commercially available, although they do not store much energy (Section 9). Nevertheless, in lab situations, and exam problems, most capacitors are microfarads (μF), nanofarads (nF) or even picofarads (pF $\equiv 10^{-12}$ F).

8. Capacitors in Series and Parallel

Like resistors, capacitors can be connected in series or parallel. However, the rules for adding the capacitors are different. Consider two capacitors, C_1 and C_2, connected across a battery in parallel, as on the left of Figure 10.8.

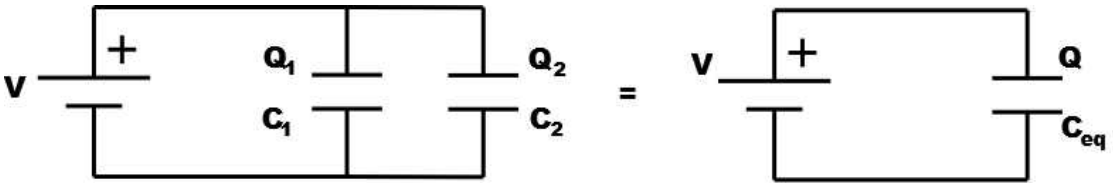


Figure 10.8

There will be a charge Q_1 on C_1 and a charge Q_2 on C_2, which are not necessarily equal, but we do know that the voltage difference across both capacitors is the battery voltage V because both capacitors are connected directly across it. Then by definition of capacitance, $V = Q_1/C_1 = Q_2/C_2$, or $Q_1 = VC_1$ and $Q_2 = VC_2$. The two capacitors together form an equivalent capacitor C_{eq} as on the right. We don't know yet what C_{eq} is, but we see that it is also connected across V. It must also carry a charge $Q = Q_1 + Q_2$. By definition $Q = C_{eq}V$, so from above, $C_{eq}V = VC_1 + VC_2$. Cancelling the V's shows that

In parallel circuits capacitances add:

$$C_{eq} = C_1 + C_2 + C_3 + \cdots + \qquad (10.9)$$

Therefore, **capacitors in parallel add like resistors in series.**

Adding capacitors in series is trickier. We have:

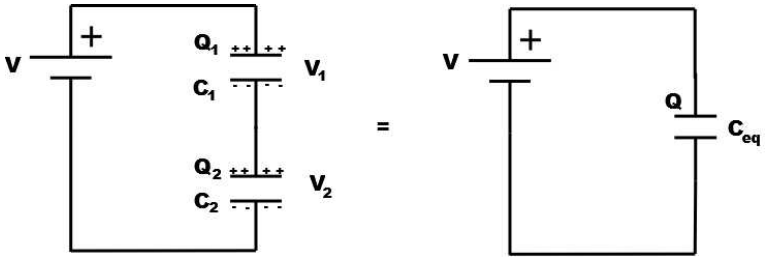


Figure 10.9

The battery on the left sends out a current from the positive terminal, but it cannot cross the gap in capacitor C_1. Instead, a positive charge Q_1 accumulates on the top plate. Q_1 induces an equal and opposite charge $-Q_1$ on the bottom plate. Because the charge between the two capacitors was initially zero, it must remain zero, which requires that an equal and opposite positive charge $+Q_1$ must appear on the top plate of C_2, implying in turn that a charge $-Q_1$ appears on the bottom plate of C_2. The magnitude of the charges on all the plates is thus the same.

Call it Q. Then $V_1 = Q/C_1$ and $V_2 = Q/C_2$. Furthermore, as with resistors in series, $V = V_1 + V_2$.

If we imagine replacing the two series capacitors with an equivalent capacitor C_{eq} such that the battery moves the same amount of charge, then the charge on the top plate of C_{eq} is also Q and the voltage across C_{eq} is the entire battery voltage V. Hence, $V = Q/C_{eq} = V_1 + V_2$ which gives from above $Q/C_{eq} = Q/C_1 + Q/C_2$. Cancelling the $Q's$ shows

In series circuits inverse capacitances add:

$$\frac{1}{C_{eq}} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} + \cdots + \qquad (10.10)$$

Therefore, **capacitors in series add like resistors in parallel.**

Warning: The same warning regarding inverse resistances applies.

Tip: The same tips regarding adding resistors in parallel apply.

Example 10.7: Typical capacitor problem. What is the equivalent capacitance of the below circuit if $C_1 = 10\ \mu$F; $C_2 = 20\ \mu$F; $C_3 = 15\ \mu$F; $C_4 = 30\ \mu$F and $C_5 = 100\ \mu$F?

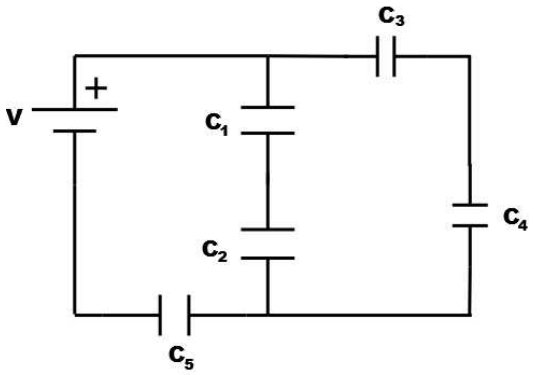


Figure 10.10

First observe that C_1 and C_2 are in series. For their combined capacitance, $1/C_{12} = 1/10 + 1/20 = 3/20$ and $C_{12} = 20/3\ \mu$F. Capacitors C_3 and C_4 are also in series. So $1/C_{34} = 1/15 + 1/30 = 3/30$ and $C_{34} = 10\ \mu$F.

Note that capacitors C_{12} and C_{34} are in parallel, and so their capacitances add directly, giving $C_{1234} = 10 + 20/3 = 50/3\ \mu$F. Lastly, C_{1234} and C_5 are in series, so the total capacitance of the circuit is given by $1/C_{eq} = 3/50 + 1/100 = 7/100$, and $C_{eq} = 100/7\ \mu$F.

Advice: Although we strongly encourage algebraic solutions, when many inverse resistances or capacitances are involved, putting terms over common denominators to solve algebraic equations can result in very cumbersome expressions and it is usually faster to plug in the numbers as one goes along.

9. Energy Stored in Capacitors

Capacitors store charge and in doing so store energy. Returning to the example of the sphere in Section 7, it takes no work to bring a small charge q from infinity to the surface of an uncharged sphere of radius R. However, once this charge resides on the sphere, it takes an increment of work equal to the change in potential energy to bring in a second bit of charge Δq to R: $\Delta U = kq\Delta q/R$. As one brings more and more charge in from infinity, it takes ever more work, because ever greater amounts of charge already reside on the sphere and repel the latest Δq. If the final amount of charge on the sphere is Q, the average charge during the entire process is $Q/2$ and the total work that has been performed is $U = kQ^2/2R$. This is the potential energy stored by a charged sphere. In Section 7 we decided that for a sphere $C = R/k$. Then $U = Q^2/2C$. This expression is valid for any capacitor and represents the energy stored in the electric field itself. Usually it is written in two equivalent ways:

$$U = \frac{1}{2}\frac{Q^2}{C} \quad (10.11a)$$

or

$$U = \frac{1}{2}CV^2 \quad (10.11b)$$

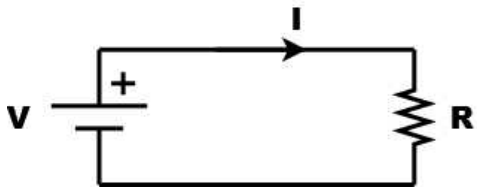
where we have used $Q = CV$ to get from the first expression to the second. As mentioned in the previous section, electrical current cannot flow through a capacitor because of the gap. Instead, charge from the battery gradually accumulates on one plate, and an equal and opposite charge accumulates on the other plate. This process is known as **charging** a capacitor. When a capacitor is fully charged, Eqs. (10.11) give the energy stored in it. If the plates of a capacitor are connected to each other through a wire, the positive charges rapidly move to the negatively charged plate until there is no separation of the charges and the net charge on each plate is zero. During this **discharge** the energy stored in the capacitor is transferred from one place to another, often accompanied by sparks. Never put your fingers across the leads of a capacitor; if the capacitor is large enough, the discharge can be lethal.

Example 10.8: Pulling plates apart. Another way to arrive at Eqs. (10.11) is to consider two parallel, oppositely charged plates. Because they are oppositely charged, the plates attract each other and it must take work to pull them apart. How much? If the charge on each plate is Q, then the force one plate exerts on the other is $F = QE$. From Eq. (9.5), the field of a large plate is $E = 2\pi k\sigma = \sigma/2\epsilon_o$. Thus, $F = Q\sigma/2\epsilon_o$ is the force that must be overcome to pull the plates apart. The work required to separate them a distance d is $W = Fd = Q\sigma d/2\epsilon_o$. But $\sigma = Q/A$. Thus, the work to separate the plates is $W = Q^2 d/2A\epsilon_o$. However, we recognize $A\epsilon_o/d$ as exactly the capacitance of a parallel-plate capacitor with a plate separation x, and so $W = Q^2/2C$, as above. Since you are doing work against the electric field, this is a strong hint that the energy stored in a capacitor is the energy stored in the field itself.

SIMPLE CIRCUITS PROBLEMS

Problems 1-20 are at the subject-test level. The remaining problems are at the AP level.

Questions 1 - 4 refer to the following figure.



1. A battery with voltage V sends a current I through a resistor R. If V is quadrupled, I

   (A) stays the same.
   (B) quadruples.
   (C) halves.
   (D) doubles.
   (E) goes down by a factor of four.

2. A battery with voltage V sends a current I through a resistor R. If R is doubled, I

   (A) stays the same.
   (B) quadruples.
   (C) halves.
   (D) doubles.
   (E) goes down by a factor of four.

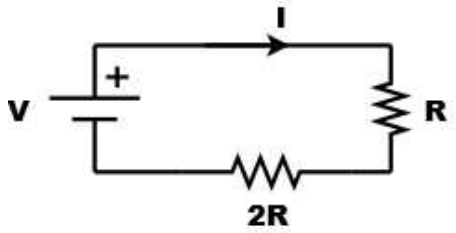
3. A battery with voltage V sends a current I through a resistor R. If V is doubled, the power dissipated by the resistor

   (A) stays the same.
   (B) quadruples.
   (C) halves.
   (D) doubles.
   (E) goes down by a factor of four.

4. A battery with voltage V sends a current I through a resistor R. If V is doubled and R halved, the power dissipated by the resistor

   (A) goes up by a factor of eight.
   (B) quadruples.
   (C) halves.
   (D) doubles.
   (E) goes down by a factor of four.

Questions 5 - 6 refer to the following figure.



5. A battery V sends a current I through two resistors in series, R and $2R$. If the first resistor is doubled to $2R$, the current I is now

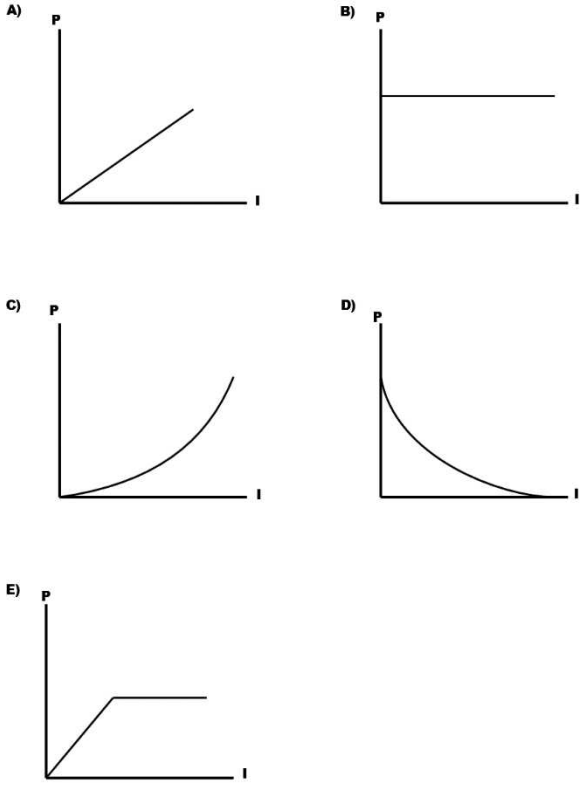
   (A) the same as it was initially.
   (B) half what it was initially.
   (C) three-quarters what it was initially.
   (D) four-thirds what it was initially.
   (E) twice what is was initially.

6. A battery V sends a current I through two resistors in series, R and $2R$. If the second resistor is halved to R, the current dissipated power

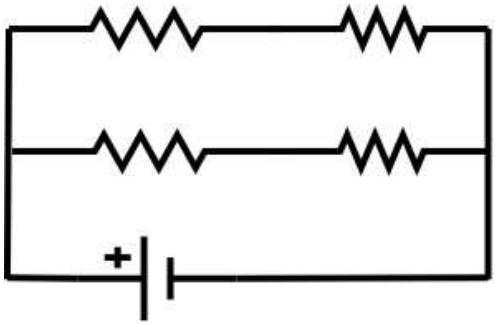
   (A) stays the same.
   (B) goes up by a factor of two.
   (C) goes down by a factor of two.
   (D) is two-thirds what it was initially.
   (E) is three-halves what it was initially.

7. Which of the below graphs best represents the power dissipated in a resistor as a function of current?

   (A) A
   (B) B
   (C) C
   (D) D
   (E) E



Questions 8 - 10 refer to the following figure.



8. Four identical resistors, each with resistance R, are arranged in the circuit above. The total resistance of the circuit is

   (A) $R/4$
   (B) $R/2$
   (C) R
   (D) $2R$
   (E) $4R$

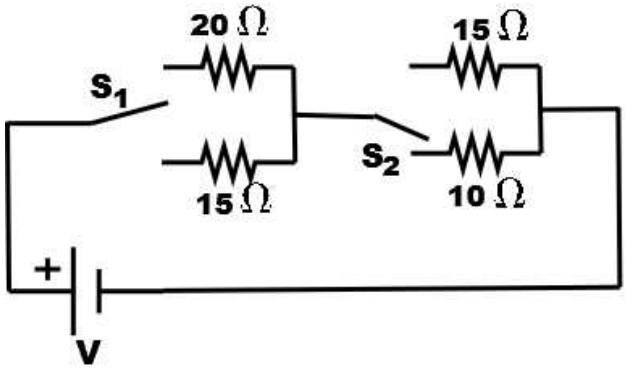
9. If the two resistors in the top leg of the circuit are each doubled in value to $2R$, the total resistance of the circuit becomes

   (A) $6R$
   (B) $4R$
   (C) $4R/3$
   (D) $3R/4$
   (E) $R/4$

10. If one resistor from each branch is doubled from R to $2R$, the total resistance of the circuit becomes

    (A) $3R$
    (B) $3R/2$
    (C) $4R/3$
    (D) $3R/4$
    (E) $2R/3$

Questions 11 - 12 refer to the following figure.



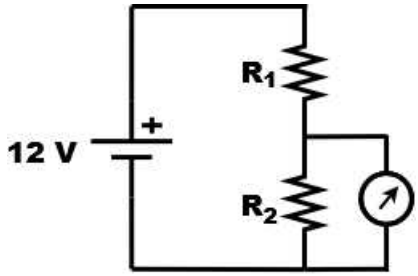
11. A circuit consists of a battery V with two switches S_1 and S_2. Each switch can be thrown either up or down, as shown. Which positions result in the greatest current through the circuit?

(A) S_1 up; S_2 up
(B) S_1 down; S_2 down
(C) S_1 up; S_2 down
(D) S_1 down; S_2 up
(E) (C) or (D)

12. A circuit consists of a battery V with two switches S_1 and S_2. Each switch can be thrown either up or down, as shown. Which positions result in the minimum current through the circuit?

(A) S_1 up; S_2 up
(B) S_1 down; S_2 down
(C) S_1 up; S_2 down
(D) S_1 down; S_2 up
(E) (C) or (D)

Questions 13 - 14 refer to the following figure.



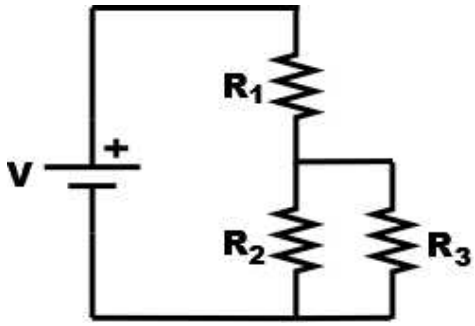
13. A battery puts out 12 V. To run an electronic device, you need 4 V. You construct the above circuit to produce 4 V across one of the resistors, R_2. Values of the resistors that would work for this purpose are

(A) $R_1 = 600\ \Omega\ ;\ R_2 = 600\ \Omega$
(B) $R_1 = 1200\ \Omega\ ;\ R_2 = 600\ \Omega$
(C) $R_1 = 600\ \Omega\ ;\ R_2 = 1200\ \Omega$
(D) $R_1 = 1200\ \Omega\ ;\ R_2 = 400\ \Omega$
(E) $R_1 = 400\ \Omega\ ;\ R_2 = 1200\ \Omega$

14. The above battery is changed to 6 V. To run the same electronic device, you could now use resistors R_1 and R_2 such that

(A) $R_1 = 600\ \Omega\ ;\ R_2 = 600\ \Omega$
(B) $R_1 = 1200\ \Omega\ ;\ R_2 = 600\ \Omega$
(C) $R_1 = 400\ \Omega\ ;\ R_2 = 1200\ \Omega$
(D) $R_1 = 1200\ \Omega\ ;\ R_2 = 400\ \Omega$
(E) $R_1 = 600\ \Omega\ ;\ R_2 = 1200\ \Omega$

Questions 15 – 17 refer to the following figure.



15. In the above circuit, $R_1 = 1\ k\Omega$; $R_2 = 1\ k\Omega$ and $R_3 = 4\ k\Omega$. The total resistance of the circuit is

    (A) 5.0 kΩ
    (B) 1.8 kΩ
    (C) 0.8 kΩ
    (D) 0.6 kΩ
    (E) 0.4 kΩ

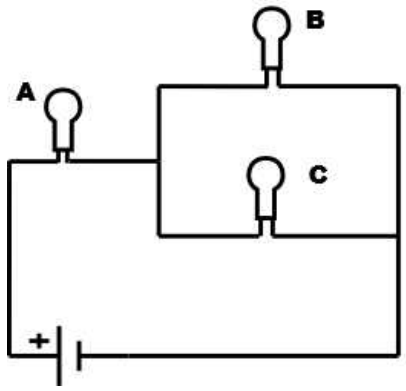
16. Assume that in the above circuit all the resistors have the same resistance R. The ratio of the power dissipated by R_3 to the power dissipated by R_1 is

    (A) $1/4$
    (B) $1/2$
    (C) 1
    (D) 2
    (E) 4

17. Assume that in the above circuit all the resistors initially have the same resistance R. The value of R_3 is halved to $0.5R$. The ratio of the final power dissipated by R_3 to the initial power dissipated by R_3 is

    (A) $8/9$
    (B) $9/8$
    (C) $4/3$
    (D) $3/4$
    (E) $2/3$

Questions 18 – 19 refer to the following figure.



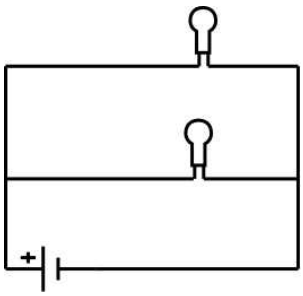
18. A circuit consists of a battery V and three identical light bulbs, each with resistance R, wired in the above configuration. Rank the bulbs in order of brightness, from brightest to dimmest.

    (A) A, B, C
    (B) C, B, A
    (C) B = C, A
    (D) A, B = C
    (E) A = B, C

19. If bulb B is removed from the circuit, the brightness of bulb C

    (A) remains the same.
    (B) goes up by a factor of four.
    (C) goes down by a factor of four.
    (D) goes down by a factor of $4/9$.
    (E) goes up by a factor of $9/4$.

20. A circuit consists of a battery V and two identical light bulbs, each with with resistance R.



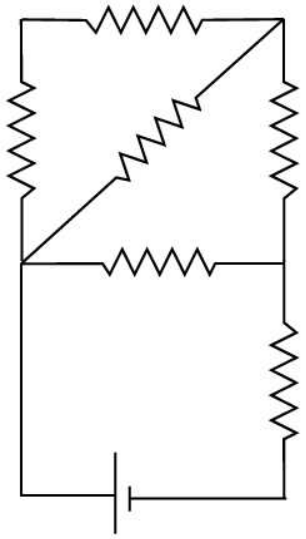
    If one light bulb burns out, the brightness of the second bulb

    (A) quadruples.
    (B) goes down by a factor of four.
    (C) stays the same.
    (D) doubles.
    (E) halves.

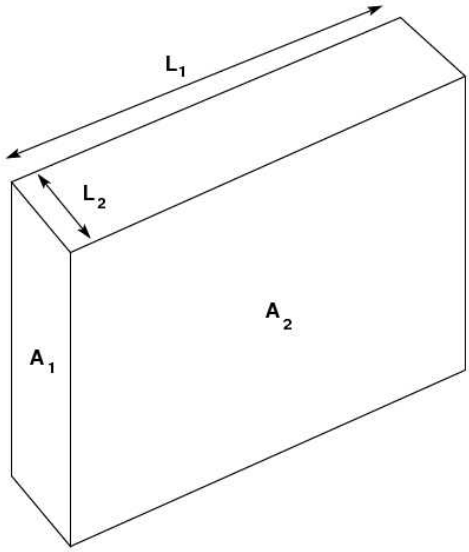
21. Six identical resistors R are connected across a battery, as shown below.

The total resistance of the circuit is

(A) $6R$
(B) $13R/8$
(C) $8R/13$
(D) $6R/27$
(E) $27R/6$



Questions 22 – 24 refer to the following figure.



22. A rectangular copper block has a resistivity 1.7×10^{-8} Ω m. $L_1 = 15$ cm; $L_2 = 3$ cm. The area of the faces are $A_1 = 18$ cm^2 and $A_2 = 90$ cm^2. A current is passed through the length L_1. The resistance of the block is

(A) 1.4×10^{-6} Ω
(B) 1.4×10^{-10} Ω
(C) 5.7×10^{-8} Ω
(D) 5.7×10^{-4} Ω
(E) 7.1×10^{-6} Ω

23. If all the dimensions of the block are doubled, the resistance

    (A) quadruples.
    (B) goes down by a factor of four.
    (C) stays the same.
    (D) doubles.
    (E) halves.

24. For a rectangular block with $L_1 > L_2$, as drawn, can the resistance with the current through the length L_2 be greater than the resistance with the current going through the length L_1?

    (A) Always
    (B) Sometimes, if L_2 is long enough
    (C) Sometimes, if A_2 is large enough
    (D) Sometimes, if A_1 is small enough
    (E) Never

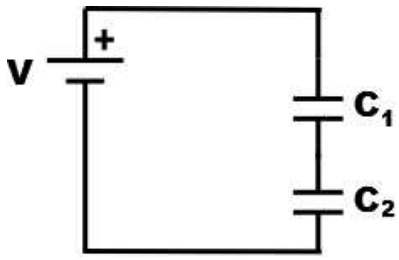
25. A camera flash releases 3 J of energy, which is supplied by a capacitor. If the voltage of the flash unit is 200 V, the capacitance of the capacitor is

    (A) 200 F
    (B) 200 μF
    (C) 150 μF
    (D) 10^{-8} F
    (E) 150 nF

26. If the energy of the flash in the previous problem is released over a millisecond, what is the average power produced?

    (A) 3000 W
    (B) 1500 W
    (C) 1000 W
    (D) 200 W
    (E) 100 W

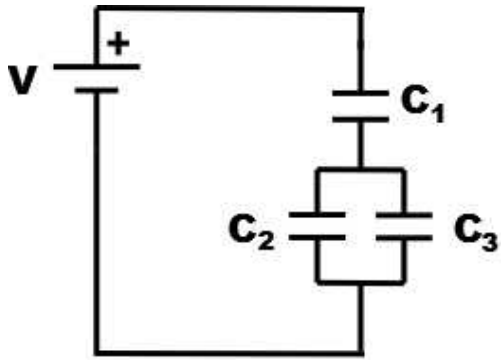
27. Two capacitors are connected in series to a voltage source, as shown below.



    If $V = 500; C_1 = 2\ \mu$F and $C_2 = 8\ \mu$F, what is the voltage drop, V_1 and V_2, across each?

    (A) $V_1 = 250$ V; $V_2 = 250$ V
    (B) $V_1 = 20$ V; $V_2 = 80$ V
    (C) $V_1 = 200$ V; $V_2 = 300$ V
    (D) $V_1 = 400$ V; $V_2 = 100$ V
    (E) $V_1 = 100$ V; $V_2 = 400$ V

Questions 28– 29 refer to the following figure.



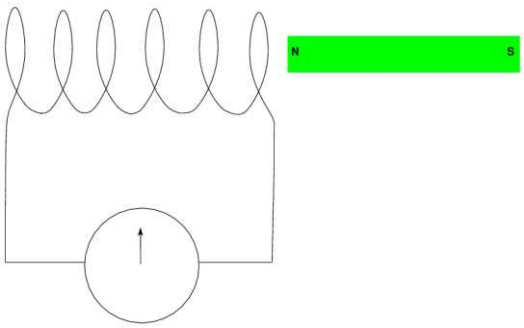
28. Three capacitors are connected to a voltage source, as shown. If $V = 500$ V; $C_1 = 4\ \mu$F; $C_2 = 2\ \mu$F and $C_3 = 10\ \mu$F, what is the voltage drop, across C_1?

(A) 75 V
(B) 125 V
(C) 200 V
(D) 275 V
(E) 375 V

29. If $V = 480$ V; $C_1 = 10\ \mu$F; $C_2 = 2\ \mu$F and $C_3 = 4\ \mu$F, rank the capacitors in terms of stored energy, from highest to lowest.

(A) $C_3, C_1 = C_2$
(B) C_3, C_1, C_2
(C) C_1, C_2, C_3
(D) $C_1 = C_3, C_2$
(E) C_2, C_3, C_1

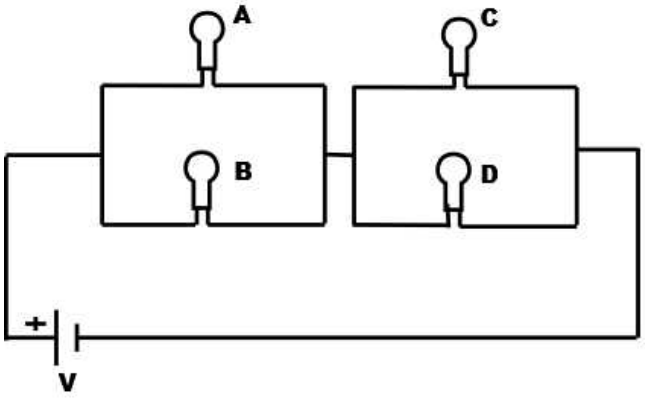
30. A bar magnet of 0.1 T is thrust into a coil of wire from a large distance in 0.5 s. If the coil has 1000 turns, an area of 0.01 m^2 and its resistance is 0.5 Ω, what is the approximate current going through the wire?



(A) 4 A
(B) 2 A
(C) 0.5 A
(D) 0.05 A
(E) 0.01 A

Free-response problems

31. A battery of 36 V is connected to four light bulbs as shown in the diagram. The resistances of the bulbs are: $A = 4\ \Omega$; $B = 12\ \Omega$; $C = 24\ \Omega$; $D = 8\ \Omega$.



a) Which direction (clockwise or counterclockwise) are electrons moving in this circuit?

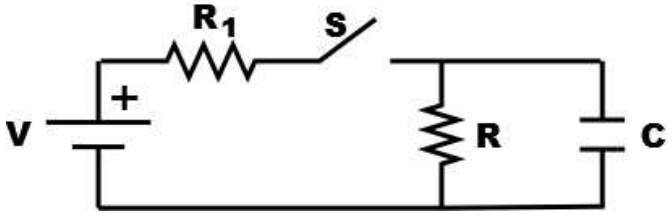
b) Are the electrons gaining or losing potential energy as they traverse the circuit?

c) What is the value of the current coming out of the battery?

d) What is the current going through light bulb C?

e) Rank the bulbs in order of brightness, from brightest to dimmest.

32. A 24-volt battery is connected to two resistors and a capacitor in the circuit shown below. The resistor marked R_1 has a value of 1 kΩ; the resistor marked R has a value of $1500\ \Omega$. The capacitor C has a value of $50\ \mu$F. Initially the switch S has been *closed* for a long time.



a) What is the value of the current through R_1?

b) Draw in an ammeter in the correct position to measure the current through R_1 and a voltmeter to measure the voltage drop across it.

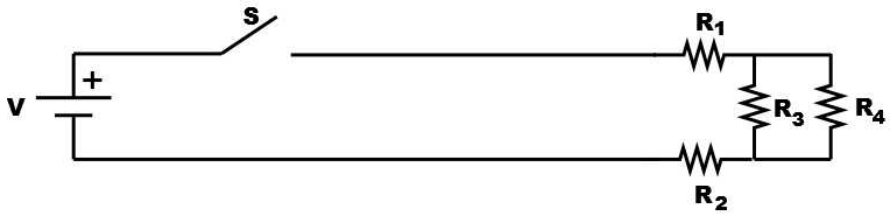
c) What is the charge on the capacitor C at this time? Energy?

d) The switch S is then opened. What happens to the charge on C?

e) What is the energy dissipated by R, if any.

f) The resistor R is replaced by a $4500\ \Omega$ resistor. S is closed again. After a long time, what is the energy on C compared to the result in (c)?

33. A 2400-volt battery is connected to itself through two long wires and four resistors. The values of the resistors are: $R_1 = R_2 = 10\ \Omega$; $R_3 = 20\ \Omega$; $R_4 = 5\ \Omega$. The top and bottom wires are 20 cm apart. The mass per unit length of the wire is 100 gm/m.



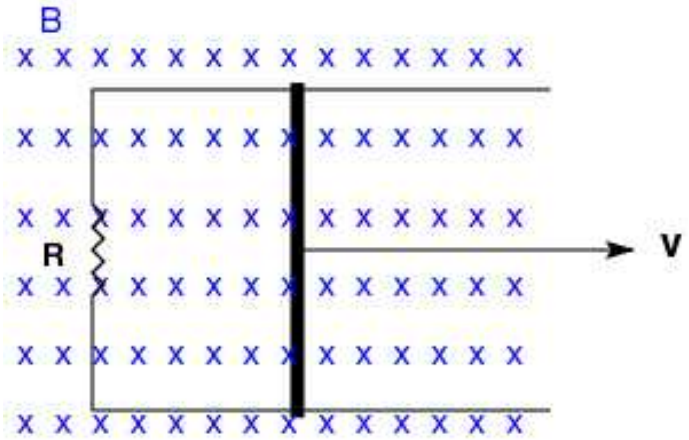
a) At a time $t = 0$ switch S is closed. What is the current I through the wires?

b) What is the direction and strength of the magnetic field produced by the top wire at the position of the bottom wire?

c) What is the direction and magnitude of the force per unit length on the bottom wire?

d) What will one observe when one closes the switch? Will this action be momentary or prolonged?

34. A bar with length $\ell = 0.4$ m and mass $m = 0.3$ kg sits on two rails, as shown. The electrical resistance of the rails themselves is negligible, but they are connected through a resistor $R = 4\ \Omega$. A magnetic field of 0.6 T passes through the rails into the page. The rail is pulled at a velocity of 1.2 m/s to the left. A coefficient of friction $\mu = 0.2$ exists between the bar and the rails.



a) What is the magnitude of the current induced in the resistor?

b) What is the force required to pull the bar at the given velocity?

c) What is the power dissipated in the resistor?

d) What is the power expended by the force on the bar in pulling it?

e) Why or why not are the answers in (c) and (d) the same?

11. OPTICS REVIEW

The material covered by this chapter appears on both the SAT subject test and the AP exams. Sections 11.1-11.10 are at the subject-test level. Sections 11.11-11.12 are at the AP level.

1. Essential Terms

The subject of **optics**, from the Greek *optikos*, "having to do with light," is indeed the study of the behavior of light. Light is a traveling electromagnetic wave. As such, it shares with all other waves the basic wave properties reviewed in Chapter 5. In particular, any light wave has a **frequency** and a **wavelength**. However, the wavelength of visible light is of order 10^{-7} m. This makes it extremely short compared to the size of most objects encountered in daily life. In such situations we can treat light as a **ray** that travels from one point to another in a straight line, unless it is reflected by a **mirror** or **refracted** by a **prism** or **lens**. In that case, the subject is termed **ray optics** or **geometric optics.** Ray optics is the focus of Sections 2-9. The SAT subject test assumes a qualitative knowledge of ray optics.

It is often the case, however, that light passes through narrow **slits** in a screen or through **thin films** of soap, oil or special lens coatings. In such scenarios, the physical dimensions of the system may be comparable to the wavelength of light. In this regime the subject of optics is termed **wave optics** because the wave nature of light must be taken into account. It is in the domain of wave optics that the phenomena of **interference** and **diffraction** of light are observed.

Light waves differ from sound waves or water waves in that they also exhibit a property known as **polarization**, which refers to the fact that light waves oscillate only in specific directions. Because light is or can be polarized, its **intensity** changes when a beam passes through certain materials, such as sunglasses, that are known as **polarizers**. A full understanding of polarization requires wave optics, indeed quantum mechanics, but the basic results can be stated in terms of rays. The SAT subject test assumes a qualitative understanding of wave optics and the AP exam expects a more quantitative understanding. Polarization is reviewed in Section 10. Wave optics is treated in Sections 11-12

2. The Law of Reflection

The basic assumption of ray optics is that **light rays travel in straight lines** until they encounter an obstacle. A ray is a collection of light waves that propagate in the same direction, behaving as a single entity. Because light rays travel in straight lines, a **ruler** or **straight edge** is an indispensable tool for problem solving, which in optics often requires more geometry than algebra.

A **mirror** is an object that reflects light such that it preserves an image. The basic rule of reflection, which may be regarded as the definition of a mirror, is the

Law of reflection:

The angle of reflection equals the angle of incidence

The law of reflection is illustrated in Figure 11.1. It says simply that a light ray bounces off a surface at the same angle with which it strikes the surface; the angle of reflection equals the angle of incidence: $\theta_r = \theta_i$. In this sense light behaves like material particles, in particular billiard balls, which observe the same law on a pool table under elastic collisions. If an object reflects light according to the law of reflection, then we term it a mirror. The important point to remember is that by convention, **all angles in ray optics are measured with respect to the normal to the surface.**

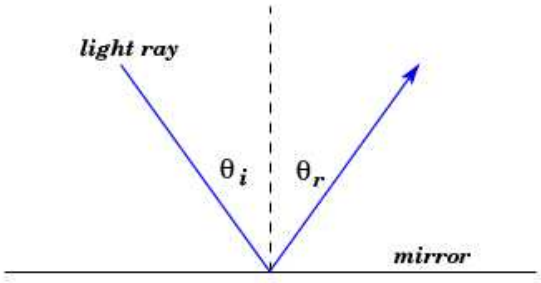


Figure 11.1

3. The Law of Refraction

When a light ray passes from one transparent medium to another, such as from air to water, we observe that the ray bends, or is **refracted**. Figure 11.2 illustrates the behavior.

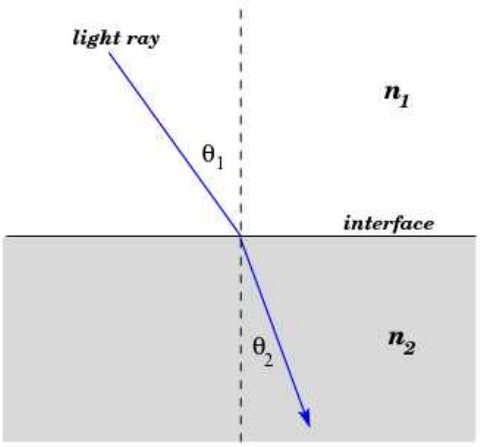


Figure 11.2

The interface can represent the boundary between air and a swimming pool or between air and a lens, etc. Each medium is characterized by a number n termed the **index of refraction** or **refractive index**. Typically, the denser the medium, the higher the index of refraction. For vacuum, $n = 1$. For air, to a high approximation, $n = 1$. For water $n = 4/3$. For most glasses $1.5 < n < 2$. Diamond has the highest index of refraction of any common material: $n = 2.42$. All these numbers are worth depositing in your number bank.

If a light ray passes from one medium with an index of refraction n_1 into a medium with a higher index of refraction n_2, the ray is bent *toward* the normal; this is the situation drawn in Figure 11.2. In that case the angle of incidence is θ_1, which is greater than the angle of refraction θ_2. On the other hand, if we retrace the path of the light ray from medium 2 to medium 1, where $n_1 < n_2$, the ray is refracted *away* from the normal. In this case θ_2 becomes the angle of incidence and θ_1 becomes the angle of refraction. (Because refraction works "both ways" we do not label the angles θ_i and θ_r.) Experiment establishes

The law of refraction:

$$n_1 \sin\theta_1 = n_2 \sin\theta_2 \quad (11.1)$$

aka **Snell's law**, after Willebrord Snellius.

Exercise: Assuming that $n_1 = 1$ and $n_2 = 4/3$, make a figure like Figure 11.2 and draw several rays that obey Snell's law. In particular, observe the behavior of θ_2 as θ_1 is made larger or smaller.

Example 11.1: Snatching at pennies. A typical situation in which refraction is important occurs when you attempt to pick up a penny from the bottom of a fountain.

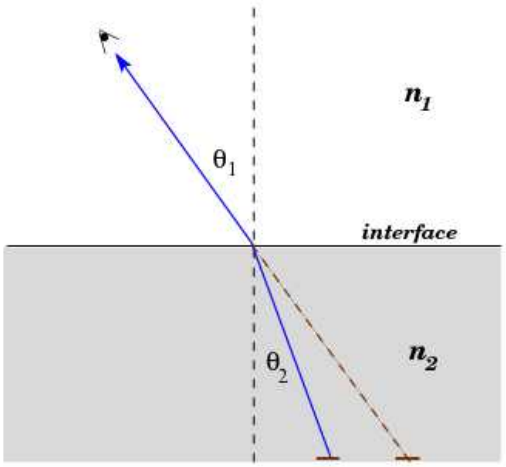


Figure 11.3

A light ray proceeds from the penny upwards at angle θ_2 through the water ($n_2 = 4/3$) until it strikes the air ($n_1 = 1$). Since $n_1 < n_2$, it is refracted at angle θ_1 away from the normal until it hits your eye. Your eye retraces the ray along a straight path (dashed line) to the bottom of the fountain, where it places the penny. Thus, you believe the penny is farther away than it really is, and you invariably fail to grab it on the first few tries.

Exercise: If x is the true distance of the penny from the indicated normal, and d is the depth of the fountain, find the apparent distance X of the penny from the normal in terms of d, x and n_2.

The law of refraction immediately gives rise to an important phenomenon known as **total internal reflection.** Suppose that in Figure 11.3 a ray is emitted by the penny at an angle θ_2 such that $\theta_1 = 90°$. In other words, the ray is refracted along the surface of the water and never gets out., as in Figure 11.4

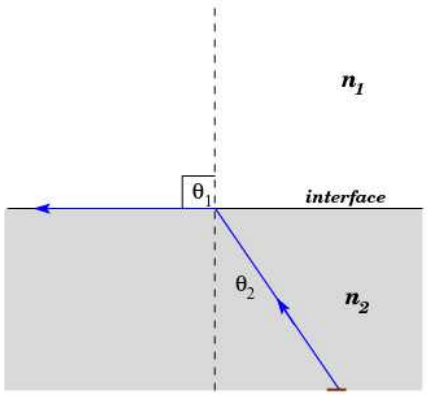


Figure 11.4

Since $\sin 90° = 1$ the law of refraction gives $n_1 = n_2 \sin\theta_2$. For angles $\theta_1 > 90°$ the law of refraction ceases to make sense, since the light is no longer refracted at all, but reflected back into medium 2. Thus, $\theta_1 = 90°$gives the maximum allowed value of θ_2, which we label θ_{crit}. Then

$$\theta_{crit} = \sin^{-1}\frac{n_1}{n_2} \quad (11.2)$$

When $\theta_2 > \theta_{crit}$ no light escapes medium 2. If medium 1 is air and medium 2 is water, then $n_1 = 1$ and $n_2 = 4/3$, giving $\theta_{crit} = 0.848$ rad, or 48.6°. Any beam of light, for instance from a laser, aimed at the surface of the water at a greater angle will be totally reflected back into the water according to the law of reflection. Fiber optics rely on total internal reflection to transmit communication signals. Prismatic binoculars would be impossible without total internal reflection.

Example 11.2: Prismatic binoculars. A **prism** is an optical element, typically made of glass or plastic in a triangular shape, that refracts light without forming an image. In the correct orientation, a prism can be used as a **corner reflector**, which reflects light back in the direction it came. Such corner reflectors have been placed on the moon in order to provide an accurate measurement of the lunar distance by reflecting laser light from earth. Corner reflectors also provide critical components in binoculars by reflecting light in the required direction. Suppose a pair of binoculars contains a prism, as in Figure 11.5a.

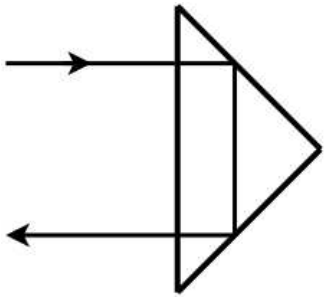


Figure 11.5a

Assume that the prism is a 45-45-90 triangle. What must be the index of refraction of the glass in order to use the prism as a corner reflector?

Drawing in the normal at the point where the light beam strikes the prism (Figure 11.5b), we see that the angle of incidence is 45°.

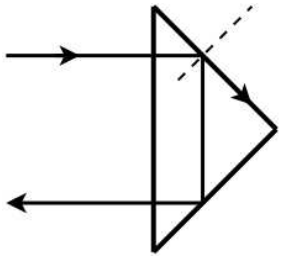


Figure 11.5b

For total internal reflection to occur, the angle of refraction θ_2 must be at least 90°, which sends the refracted ray along the surface of the prism, as drawn. If $n_2 = 1$ for air, Snell's law then gives $n_1 \sin 45° \geq 1$. But $\sin 45° = 1/\sqrt{2}$, and so we require $n_1 \geq \sqrt{2} \approx 1.41$.

4. Refraction and Wavelength

The early investigators of refraction regarded the refractive index merely as the constant ratio of the sines of the two angles in Eq. (11.1). That is, if $n_1 = 1$, then $n_2 = \sin\theta_1 / \sin\theta_2$. This is in fact how you determine indexes of refraction. Today, however, we know that that the law of refraction results from the fact that **the index of refraction is the ratio of the speed of light in a vacuum c to the speed of light in the medium v:**

$$n = \frac{c}{v} \quad (11.3)$$

The fact that $n = 4/3$ for water means that the speed of light in water is three-quarters the speed of light in vacuum. **Light entering material bodies slows down.**

The fundamental relationship between frequency and wavelength (Eq. 5.15) tells us that the speed of a wave is equal to its frequency multiplied by its wavelength, $v = \lambda f$. If light slows when entering a medium with refractive index $n > 1$, then either its wavelength or frequency must decrease, or both. However, frequency is a measure of the number of wave crests passing an observation point per unit time. If f decreased at the interface between the vacuum and glass, then the number of waves passing the interface per second would suddenly go down, causing waves to "pile up." Equivalently one can say that the energy transmitted by waves would be unequal on the two sides of the interface, in violation of energy conservation. Thus,

At the interface between two media, the frequency of a wave remains the same, but the wavelength changes.

Because the frequency is unchanged at the boundary, $c = \lambda_o f$ for vacuum and $v = \lambda_n f$ for the medium. Hence, Eq. (11.3) can be rewritten as

$$n = \frac{\lambda_o}{\lambda_n} \quad (11.4)$$

In other words, the index of refraction is the ratio of the wavelength of light in the vacuum to the wavelength in the medium under consideration. We will use this relationship in Section 12. Note that when light enters water, its wavelength shrinks to only three-quarters its wavelength in vacuum.

Furthermore, if you have ever shone sunlight through a glass prism, an experiment first performed by Newton, you have observed the phenomenon for which prisms are best known: a rainbow of colors emerges from the other side (Figure 11.6).

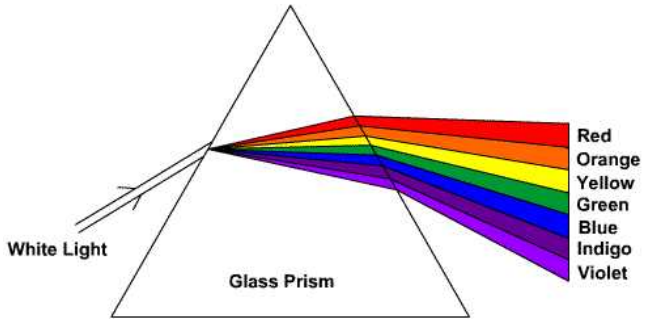


Figure 11.6

That a prism can separate white light into a **spectrum** of colors is due to two facts: One, white light is composed of many different wavelengths, each of which is associated with a different color. Two, the refractive index of a material usually depends on wavelength. Therefore, the angle of refraction for different colors varies, allowing a prism to separate them. Historically this ability made prisms the crucial element in **spectrometers**, instruments used to measure the spectra of flames and starlight, which proved the key to understanding atomic and stellar structure. The variation of refractive index with wavelength is referred to as **dispersion**, and any medium with an index of refraction that depends on wavelength is termed a **dispersive medium**.

Example 11.3: Basic spectrometer. White light is incident at 60° on a glass prism in the form of an equilateral triangle. The refractive index for red light is $n_r = 1.513$ and the refractive index for violet light is $n_v = 1.532$. At what angles do the red and violet rays exit the prism?

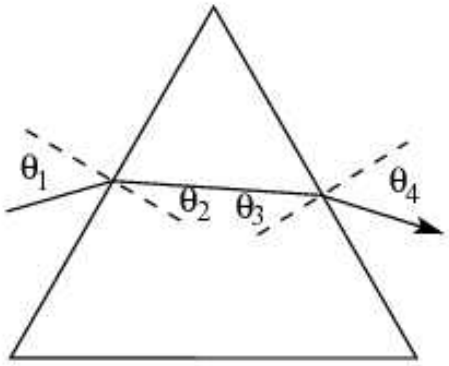


Figure 11.7

We need to find θ_4, the exit angle. The white light is incident at $60°$, meaning as always $60°$ from the normal. Therefore, $\theta_1 = 60°$ in Figure 11.7. The ray is then refracted at angle θ_2. Assuming that $n_1 = 1$ in air, Snell's law gives $\sin 60° = 1.513 \sin \theta_2$ for the red ray. Solving for θ_2 yields $\theta_2 = 34.92°$. The refracted ray travels across the prism, making an angle θ_3 with the normal on the opposite side. **Exercise:** By considering the small, upper triangle formed by the two sides of the big triangle and the ray itself, show that $\theta_3 = 60 - \theta_2$. Thus, $\theta_3 = 25.08°$. The law of refraction then gives $1.513 \sin 25.08 = \sin \theta_4$. Solving for θ_4 yields $\theta_4 = 39.89°$.

Repeating the calculation with n_v gives $\theta_4 = 41.41°$. Notice that this is larger than the value for red light. **Short wavelengths are refracted more than long wavelengths.** (You may have noticed that θ_2 for violet is *smaller* than θ_2 for red, but this means that violet is bent *more* toward the normal, consistent with the previous statement.)

Often the overall amount of refraction in a prism is given by the *deviation angle*, Δ, which is the angle that the ray is deflected from its original direction. **Exercise:** Show that $\Delta = \theta_1 + \theta_4 - 60°$.

5. Converging Lenses

You have probably had many experiences with lenses in your lifetime, either in cameras or as magnifiers. A **lens** is an optical device, typically made of glass or plastic, that transmits light, refracting it in the process. Unlike prisms, which do not form images, the usual purpose of a lens is precisely image formation. Lenses can be complicated objects, and the equations governing them as well. We will restrict ourselves to the simplest situations.

The most familiar type of lens is the **converging lens** (Figure 11.8). A converging lens consists of two convex surfaces, which take parallel rays of light and bring them together by refraction at a single point, termed the **focus** or **focal point**. Converging lenses serve as magnifying glasses and the main elements in camera lenses, telescopes, microscopes and in eyeglasses for farsighted individuals.

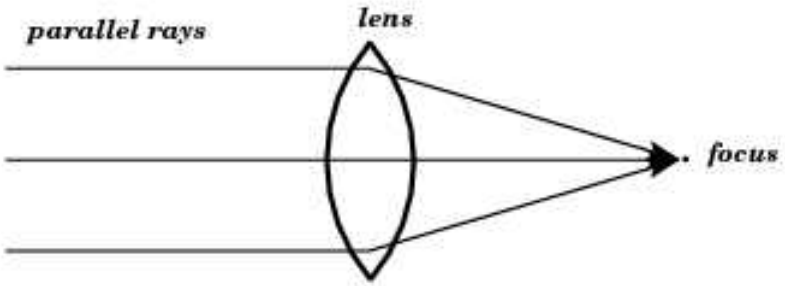


Figure 11.8

A converging lens focuses parallel rays to a point. Conversely, if rays emanate from the focus and strike the lens they emerge from the lens as a parallel, or **collimated**, light beam. Moreover, if the surfaces of the lens are symmetric, the lens works the same in both directions: A beam of parallel rays striking the lens from the left is focused on the right; a parallel beam of rays striking the lens from the right is focused on the left. Thus, **a simple lens has two focal points, located equidistantly to the left and right of the lens center**.

The emphasis on parallel rays is not arbitrary. Any light source, for instance the sun, emits rays that diverge from each other at many different angles. However, as the distance to the source grows larger, the angle between any two rays that enter your eye or a lens gets smaller and smaller. As the distance to the source becomes infinite, the angle becomes zero, which implies that the rays are parallel. **Light from infinitely distant sources consists of parallel rays.** Light rays from astronomical objects can always be taken to be parallel.

The basic process of image formation by lenses can be understood by the use of **ray diagrams**.

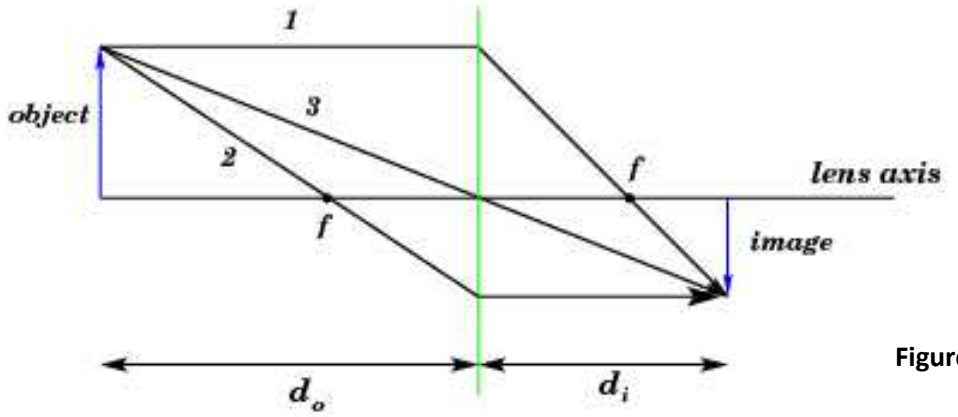


Figure 11.9

Consider an object (traditionally represented by an arrow) placed at a certain distance from the lens, called the **object distance** d_o. The lens forms an image at a certain distance from the lens called the **image distance** d_i. We assume that the lens is much thinner than d_o, d_i or the **focal length** f. Such a lens is imaginatively called a **thin lens** and can be represented by a straight line. To determine graphically where the image is formed, we use the method of "principal rays." Any rays passing through the lens can be used to construct the image, but most require making accurate measurements of angles. The "principal rays," however, are particularly simple to employ, largely because of the basic lens properties you have digested above.

1. Any ray from the object traveling parallel to the lens axis, is refracted by the lens to the far focus. Therefore, **draw a ray from the tip of the object parallel to the lens axis to the lens and then through the far focus.**

2. Any ray that passes through a focus and strikes the lens emerges from the lens on the far side parallel to the lens axis. Therefore, **draw a ray through the near focus to the lens, then parallel to the lens axis on the opposite side of the lens.**

3. Any ray passing directly through the center of the lens is unrefracted. This follows from symmetry: If a ray in Figure 11.9 strikes the surface of the lens in such a way that it is bent through the center, then as it exits it will be refracted in exactly the opposite amount, such that the net result is zero. Therefore, **draw a lens from the tip of the arrow through the center of the lens and to the opposite side.**

The point at which the drawn rays converge will be the tip of the image. In principle, one can make do with two rays, but it is better to use three.

Tip: Small drawing errors in position and direction become magnified. Use a straightedge.

In the case illustrated in Figure 11.9, the image is **inverted**. It is also termed a **real image** because if you place a screen at the plane where the image has been formed, light rays are actually going through the screen. **If the object is outside the focal length of a converging lens, the image will be real and inverted.**

An accurately made ray diagram can show *without calculation* the qualitative relationships between object and image, as well whether an image is real or virtual. For that reason, you should *always* draw one when attempting an optics problem. A ray diagram and basic geometry also allow us to deduce the algebraic relationships between the principal quantities.

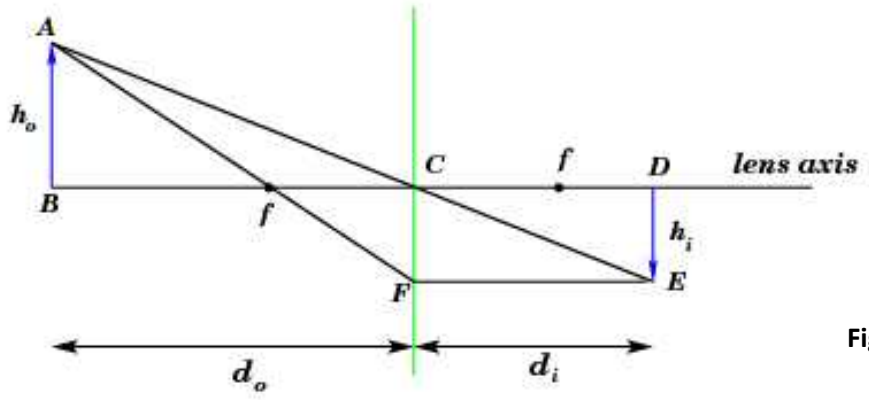


Figure 11.10

Consider Figure 11.10, where we have let h_o be the object height and h_i be the image height. From the figure we see that triangles ABC and CDE are similar. Therefore,

$$\frac{h_o}{h_i} = \frac{d_o}{d_i} \quad (11.5)$$

Triangles ABf and fCF are also similar. Therefore,

$$\frac{h_o}{h_i} = \frac{d_o - f}{f}.$$

Substituting d_o/d_i for h_o/h_i into this equation and dividing by d_o gives the

Thin lens equation:

$$\frac{1}{f} = \frac{1}{d_o} + \frac{1}{d_i} \quad (11.6)$$

which forms the basis of most elementary optics problems. The important point to notice about the equation is that because the focal length f is a constant fixed by the properties of the lens, if d_o gets smaller, then d_i must get larger, and vice versa. In particular, note that if an object is very far away such that $d_o \approx \infty$, then $d_i \approx f$. Thus, for distant objects the image forms very near the focal plane of the lens, which is the usual case for camera pictures. If, on the other hand, $d_o = f$, then $1/d_i = 0$, which implies that $d_i = \infty$. Physically, this means that no image is actually formed. Remember that rays emerging from the focal point form a parallel beam, which is not an image.

Exercise: What is the image distance d_i in terms of f if the object distance $d_o = 10f,\ 8f,\ 6f,\ 4f,\ 3f$ and $2f$?

As shown by Eq. (11.5), the image size and the object size are generally different. The **magnification** of a lens is defined as

$$m \equiv \frac{h_i}{h_o} = -\frac{d_i}{d_o} \quad (11.7)$$

The important message of Eqs. (11.5) and (11.7) is that **the ratio of the object size to the image size equals the ratio of the object distance to the image distance**. We will explain below the mysterious negative sign inserted before the d_i/d_o. Nevertheless, you can see that if the image is smaller than the object, then $m < 1$ and vice versa. For applications such as a magnifying glass, one naturally wants $m > 1$.

Example 11.4: Magnifying glass. Suppose that the object is placed inside the focal length of a converging lens. Let us find the size and position of the image.

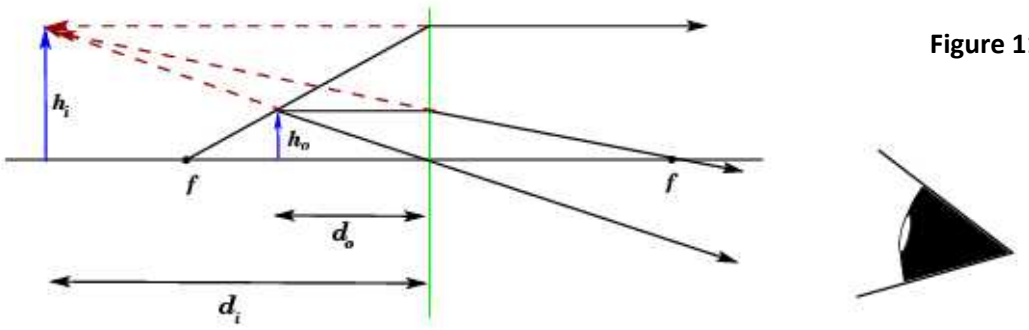


Figure 11.11

Figure 11.11 illustrates. The object of height h_o is placed *within* the focal length of the lens. We draw the three principal rays from the tip of the object to the right, as was done in Figure 11.9. (**Exercise:** Draw these rays on a separate sheet of paper and understand why they are drawn as shown.)

However, unlike the case in Figure 11.9, we see that the rays are **diverging** to the right and form no real image anywhere. The eye retraces the rays (dashed lines) until they converge at a point to the left of the lens—on the same side of the lens where the object is located. The image so formed is termed a **virtual image**, because no light rays are actually passing through any screen placed at the image's position. Rather, your brain is creating "virtual rays," which converge to form a virtual image. This result suggests a moral:

Convex lenses create either real or virtual images depending on the position of the object. If the object is outside the focal length, the image will be real and inverted. If the object is inside the focal length, the image will be virtual and upright.

Reality check: The test for a real image is to put a screen at the image's position. If you see an image on the screen it is a real image; if not, it is a virtual image.

Figure 11.11 shows that a magnifier creates an upright virtual image with $h_i > h_o$. Hence, the magnification m is greater than 1, which is what you want from a magnifier. However, suppose that $f = 15$ cm and $d_o = 10$ cm. (We have already stipulated that the object is inside the focal point of the lens). The thin lens equation, Eq. (11.6) then requires $1/15 = 1/10 + 1/d_i$. Since $1/10 > 1/15$, this implies that d_i is *negative*. Precisely, $d_i = -30$ cm.

If you are confused, you are not the first. Such a state of affairs can be understood as follows. The pair of right triangles in Figure 11.11 with heights h_i and h_o and bases d_i and d_o are similar, which immediately implies that $h_i/h_o = d_i/d_o$, as already found in Eq. (11.5). The other pair of right triangles with heights h_i and h_o and bases f and $f - d_o$ are also similar (**Exercise:** show this), giving

$$\frac{h_o}{h_i} = \frac{f - d_o}{f}.$$

Eliminating h_o/h_i in favor of d_o/d_i as before and dividing by d_o yields

$$\frac{1}{f} = \frac{1}{d_o} - \frac{1}{d_i}.$$

This expression can be reconciled with the thin lens equation (11.6) if we here define d_i to be negative. This is the first of a long list of confusing sign conventions employed in optics. We will summarize them in Section 9. For now, we state that as long as the light is coming from the side on which the object is located, the object distance d_o can be taken to be positive. If the image is formed on the opposite side of the lens from the object, the image distance d_i is also positive, as was assumed in deriving the thin lens equation. *If the image is on the same side of the lens as the object, as in this example, the image distance d_i is taken to be negative*. The motivation for the sign conventions is to be able to use Eq. (11.6) consistently for every situation involving lenses and mirrors. The signs tell you what side of the lens the image is on and whether it is upright or inverted.

This last remark pertains to the mysterious minus sign in Eq. (11.7). For the magnifier, d_i is negative, which makes $-d_i/d_o$ positive and the magnification m positive ($m = +2$). However, in the example of the converging lens with the object outside the focus, h_i/h_o is evidently positive but $-d_i/d_o$ is evidently negative (since both d_o and d_i are positive). The solution to this contradiction is to *define h_i as positive for upright images and negative for inverted images*. Thus, because the converging lens produced an inverted image, h_i/h_o becomes negative and the contradiction is removed.

As mentioned, these sign conventions are cumbersome and confusing.

Advice: If you cannot remember the sign conventions, draw an accurate ray diagram, which will tell you whether the image is real or virtual and whether it is upright or inverted.

6. Diverging Lenses

As its name implies, a **diverging lens** is one that causes light rays to diverge rather than converge (Figure 11.12). The most common application of diverging lenses is in eyeglass correction for myopia. The distinguishing feature of diverging lenses is that they consist of two concave surfaces, in contrast to the convex surfaces of a converging lens.

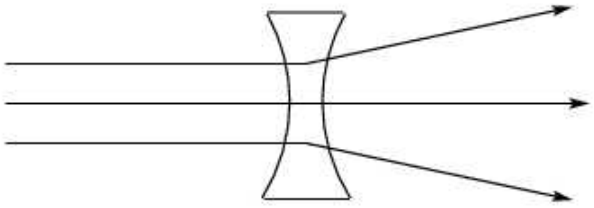


Figure 11.12

We can investigate image formation for diverging lenses by the principal ray technique; however, it is a bit trickier than for the converging lens.

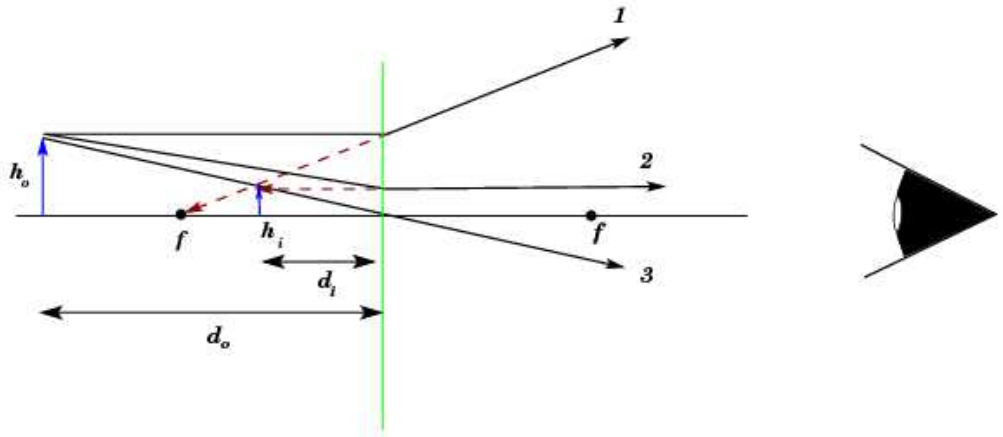


Figure 11.13

In Figure 11.13, we draw a ray (1) from the tip of the object parallel to the lens axis. Unlike the case of the converging lens, it is *not* refracted through the far focus. Instead, it is refracted *away* from the lens axis in such a way that if the eye retraces it backwards (dashed arrow) it intersects the near focus. The second ray (2) proceeds towards the far focus but is intercepted by the lens and emerges parallel to the lens axis. The eye retraces it directly backwards. The third ray (3), as before, goes straight through the center of the lens. The retraced rays converge at a *virtual, upright image* on the left side of the lens, inside the focal point.

From the similar triangles in Figure 11.13 (**Exercise**) we find

$$-\frac{1}{f}=\frac{1}{d_o}-\frac{1}{d_i}.$$

Assuming that we want d_o to remain positive, we can reconcile this equation with Eq. (11.6) if for a diverging lens we define d_i as negative and f as negative. **The focal length of a diverging lens is taken to be negative. The image distance for a diverging lens is always negative.**

Exercise: Construct diagrams similar to Figure 11.13 with the object in various positions, in particular inside the focus. Find the position and height of the image. Convince yourself that for diverging lenses:

The image is always on the object side of the lens.
The image is always smaller than the object.
The image is always upright.
The image is always virtual.

Example 11.5: Eyeglasses. In a nearsighted person, the eye's cornea and lens focus light too strongly, bringing the rays together at a point in front of the retina (Figure 11.14a), which results in a blurry image. To correct this myopia, one places a diverging lens in front of the eye (Figure 11.14b).

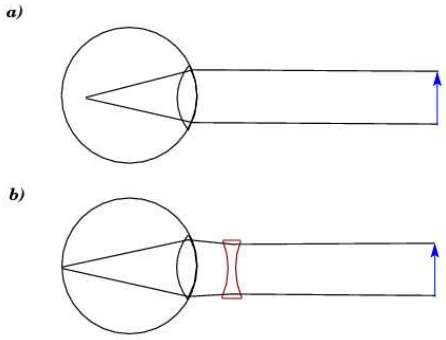


Figure 11.14

The rays from the object are spread slightly apart. The eye's lens cannot converge the spreading rays as strongly as before and so the focal point is farther away—at the retina, where the image is sharp.

Exercise: In farsighted people, the eye's cornea and lens focuses light from a distant object *beyond* the retina. What is the correction? Draw a diagram like the above.

7. Concave Mirrors

Mirrors can be employed much in the way of lenses to focus light and create images. You might think that mirrors would be simpler for this purpose, because light is merely reflected rather than refracted, but this turns out not to be entirely true. Although plane wall mirrors form images, to focus light a mirror must be curved. In real life precision mirrors (in particular telescope mirrors) are **parabolic** because parabolic surfaces bring light rays to a perfect focus. In physics courses mirrors are **spherical** because they are easier to draw. The use of spherical mirrors, though, introduces several conceptual difficulties, in particular that spherical mirrors have no well-defined focus. Incoming light rays reflected off different parts of the mirror are not brought to a single point, as shown by Figure 11.15a. This defect is known as **spherical aberration.** Nevertheless, as long as the mirror isn't too wide, it will approximate a parabolic mirror, and elementary courses always use them for demonstrations. Due to spherical aberration, however, accurate ray diagrams are difficult to draw when the object is large compared to the mirror, and in general ray diagrams for mirrors are no simpler than for lenses, if for no other reason that they require a compass or equivalent app. In the real world, the main advantages of mirrors over lenses are ease of fabrication, lower weight and cost.

Any curved mirror is concave on one side and convex on the other. The term **concave mirror** refers to a mirror whose reflective surface is on the concave side.

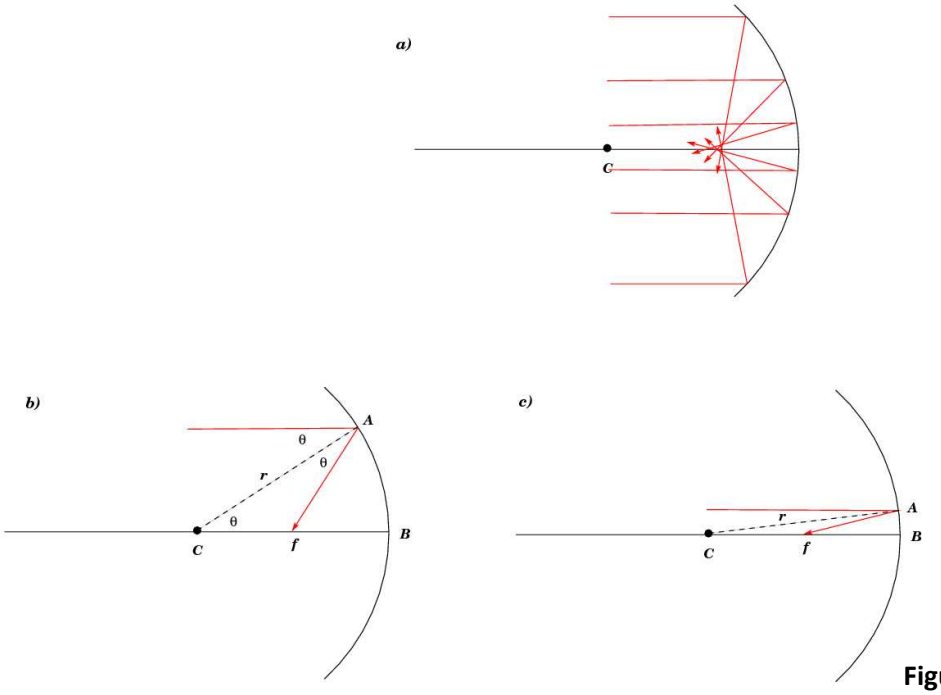


Figure 11.15

As shown in Figure 11.15b, the center of the sphere of which the mirror forms a section is marked C. It is usually referred to as the **center of curvature** of the mirror. The radius of the sphere, or **radius of curvature** of the mirror, is marked r. The radius vector from the center of the sphere to any point on the surface is perpendicular to the surface. Therefore, the dashed line CA forms the normal at point A. An incoming light ray parallel to the mirror axis strikes the mirror at A at an angle θ. By the law of reflection (Section 2) it must be reflected at an angle θ from the normal. Furthermore, by alternate interior angles, angle ACf equals θ. Triangle ACf is therefore isosceles and $Cf = fA$.

From Figure 11.15c you can see that if the ray strikes the mirror close to the axis, so that θ is small, $fA \approx fB$. Then $Cf \approx fB$ and in this limit $r = 2fB$. The distance fB is the distance to the focal point, always just called f, and so for a small spherical mirror $r = 2f$.[1]

The focus of a concave mirror is at a distance $r/2$ from the surface.

Figure 11.16 shows a ray diagram for an upright object outside the focus of a concave mirror. The object height is h_o and object distance is d_o.

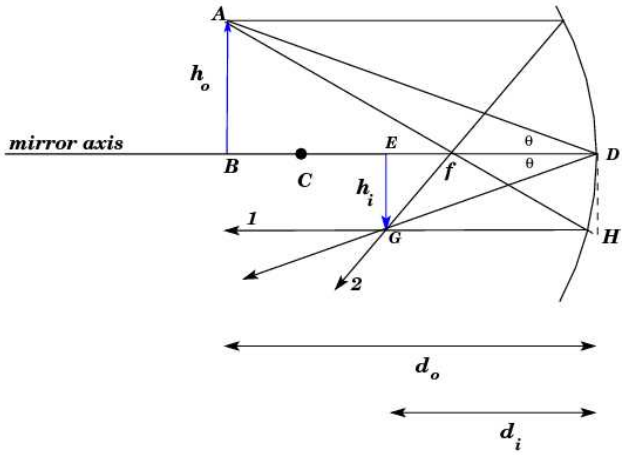


Figure 11.16

To find the image height and distance, h_i and d_i, we use the principal ray method, although the meaning of the rays is slightly different than it was for lenses:

1. A ray passing through the focus of a concave mirror is reflected outward parallel to the mirror axis. **Therefore, draw a ray from the tip of the object through the focus to the mirror and out again parallel to the mirror axis.**

2. An incoming ray parallel to the mirror axis will be reflected through the focus and out again. **Therefore, draw a ray from the tip of the object parallel to the axis until it hits the mirror, then out through the focus.**

3. A ray passing through the center of curvature of the mirror, will strike the mirror and be reflected back upon itself. We have not drawn such a ray in Figure 11.16. If the mirror or object is too large, such a ray might not intersect the others, due to spherical aberration. **For small objects, draw such a ray from the tip of the object through the center of curvature.**

In Figure 11.16, we have also drawn another ray, striking point D at an angle θ from the normal. It is therefore reflected at the same angle. This isn't usually referred to as a principal ray but one may draw any ray as long as it obeys the law of reflection. The drawn ray must also intersect the others at G. In addition, it is necessary for the following.

Looking carefully at the figure you can see that triangles ABD and DEG are similar. Therefore,

$$\frac{h_o}{h_i} = \frac{d_o}{d_i},$$

[1] An algebraic way of seeing this is by the law of cosines. Let $Cf = a$. Since triangle ACf is isosceles, the law of cosines gives $r^2 = 2a^2 - 2a^2\cos(180 - 2\theta) = 2a^2 + 2a^2\cos 2\theta$. In the limit $\theta \to 0, \cos 2\theta \to 1$ and therefore $r = 2a$.

exactly as in the previous examples. One can also see that triangle ABf is similar to triangle DfH; hence,

$$\frac{h_o}{h_i} = \frac{d_o - f}{f}.$$

Substituting d_o/d_i for h_o/h_i as before, we find

$$\frac{1}{f} = \frac{1}{d_o} + \frac{1}{d_i} \quad (11.8)$$

which we recognize as the thin lens equation. That should perhaps come as no surprise because in Figure 11.16 the *concave mirror produced an inverted real image,* just like the converging lens with the object outside the focus. You should **mentally associate converging lenses and concave mirrors** because they behave in the same way. However, remember that **for concave mirrors, d_o and d_i are positive when the object and image are in front of the mirror.**

Example 11.6: Object inside focal length. We create a ray diagram from the three principal rays discussed above.

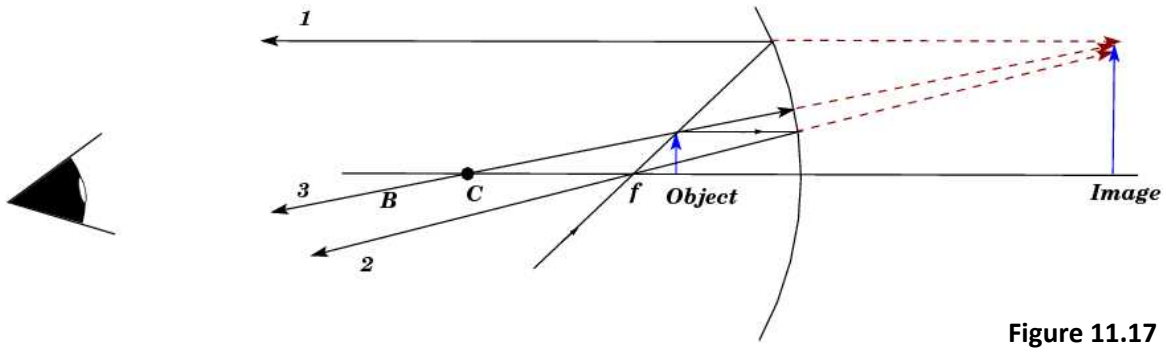


Figure 11.17

The rays are: (1), a ray that passes through the focus f, touches the tip of the object, strikes the mirror and is reflected outwardly, parallel to the mirror axis; (2), a ray from the tip of the object that travels parallel to the axis, strikes the mirror and proceeds outwardly through the focus f; (3) a ray that proceeds through the center of curvature C, touches that tip of the object and strikes the mirror, reflecting back upon itself. Once again we see that these are diverging to the left of the mirror. Your eye retraces the rays to form an upright virtual image on the right side of the mirror.

Again, we see that a concave mirror behaves like a converging lens; **when an object is within the focal length of a concave mirror, the result is an upright virtual image.** In this case the image distance d_i is also *negative*.

Exercise: Draw several ray diagrams similar to Figure 11.17 with the object in different positions.

8. Convex Mirrors

A **convex mirror** is one in which the reflective surface is on the convex side. Convex mirrors provide wide fields of view and are commonly found on cars, at the corners of buildings and at amusement parks. Just as concave mirrors are the counterpart to converging lenses, convex mirrors are the counterpart to diverging lenses. The images they create share the same properties:

The image is always behind the mirror.
The image is always smaller than the object.
The image is always upright.
The image is always virtual.

We demonstrate these properties with one ray diagram.

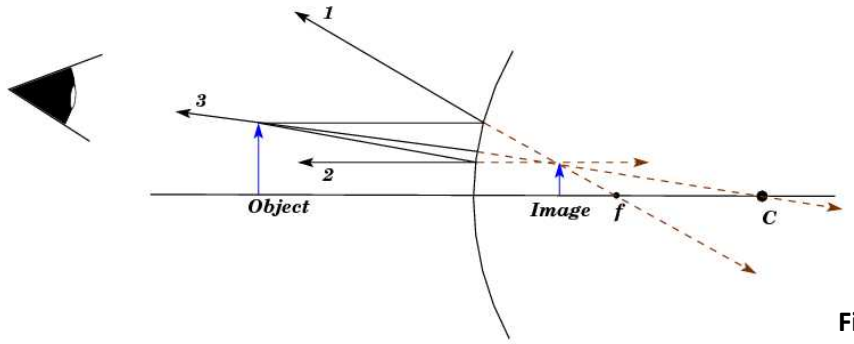


Figure 11.18

The rays used to generate Figure 11.18 are essentially those used to generate Figure 11.13 for a diverging lens. Ray 1 is drawn from the tip of the object parallel to the mirror axis. It strikes the mirror such that if retraced (dashed line) it proceeds through the focus f of the mirror. Ray 2 proceeds from the tip of the object toward the focus but is intercepted by the mirror, so it is reflected back, parallel to the mirror axis. It is retraced as a straight line parallel to the axis on the far side of the mirror. Ray 3 proceeds from the tip of the arrow toward the mirror's center of curvature C and so it is reflected back on itself; its extension passes directly through C on the far side.

We see that the image is a virtual, upright image, whose size is smaller than the object size, all in conformity with the properties listed above. You should **mentally associate diverging lenses and convex mirrors.**

Sign convention: For convex mirrors the image distance d_i and the focus are taken as negative.

Exercise: Vary the object distance and confirm that the image properties of convex mirrors always conform to the list.

Exercise: By considering similar triangles, show that the lens equation holds if the above sign conventions are observed. (*Hint:* consider the extra ray drawn in Figure 11.16.)

Example 11.7: Side view mirrors. The side view mirrors on most modern cars are convex mirrors. Suppose a Mack truck is 20 meters behind the mirror. The radius of curvature of the mirror is 18 m. What is the image distance and magnification of the mirror?

The focal length of a convex mirror is $-r/2$, where the negative sign is the convention for this mirror. Thus, $f = -9$ m. The object distance is $d_o = 20$ m. Then

$$-\frac{1}{9} = \frac{1}{d_i} + \frac{1}{20}$$

or

$$-\frac{1}{d_i} = \frac{1}{20} + \frac{1}{9} = \frac{29}{180} = .161\ m^{-1}$$

Thus, $d_i = -6.2$ m, which indicates that the image is behind the mirror. The magnification is $m = -d_i/d_o = +0.31$.

Exercise: What does the warning "Objects in this mirror are closer than they appear," mean? Does this contradict the result that $d_i < d_o$?

9. Summary of Sign Conventions for Lenses and Mirrors

	d_o	d_i	h_o	h_i	f
Convex lens (Converging lens)	+	+ for real image* − for virtual image**	+	− for inverted image* + for upright image**	+
Concave mirror (Converging mirror)	+	+ for real image† − for virtual image‡	+	− for inverted image† + for upright image‡	+
Concave lens (Diverging lens)	+	− All images virtual	+	+ All images upright	−
Convex mirror (Diverging mirror)	+	− All images virtual	+	+ All images upright	−

Notes:

d_o = object distance
d_i = image distance
h_o = object size
h_i =image size
f =focal length

The object distance d_o can always be taken to be $(+)$ if the light is coming from the same side of the lens on which the object is located. In some situations (in particular with compound lenses), the object can be on the opposite side of the lens from the light source, in which case one has a *virtual object* and $d_o < 0$.

The object height h_o can always be taken to be $(+)$.

* $d_o > f$; inverted real image on side of lens opposite object

** $d_o < f$; upright virtual image on object side of lens

† $d_o > f$; inverted real image on same side of mirror as object

‡$d_o < f$; upright virtual image behind mirror

The general rule is $(+)$ for real, upright; $(-)$ for virtual, inverted.

10. Polarization and Electromagnetic Waves

If you have ever rotated your sunglasses while looking through them at a car windshield or swimming pool, you may have noticed that the brightness of the light changes as you rotate the glasses. In making this observation, you are repeating an experiment that dates back two hundred years. As early as 1690, Christiaan Huygens passed a beam of light through a particular crystal of calcite known as Iceland spar and observed that he could extinguish the beam if he passed the light through a second crystal of Iceland spar and rotated it by the proper amount. In doing so Huygens had discovered the phenomenon of **polarization**. In 1808 Etienne Louis Malus discovered that light reflected from a window could be extinguished by rotating a crystal of Iceland spar. He went on to experimentally determine precisely how the brightness, or intensity, of the reflected light varied as the crystal was rotated:

$$I = I_o \cos^2 \theta \quad (11.9)$$

The expression, known as **Malus' law**, showed that the light's intensity I went from a maximum value I_o when the crystal was aligned in a certain position ($\theta = 0$), to zero when the crystal was rotated by an angle $\theta = 90°$ from the initial position. It is this experiment that you are confirming when you rotate your polaroid sunglasses while looking through them at a swimming pool.

The term **polarization** refers to the fact that light consists of traveling **electromagnetic waves.** Specifically, a light wave consists of an electric field E and a magnetic field B oscillating at right angles to one another, as well as at right angles to the direction of propagation (Figure 11.19).

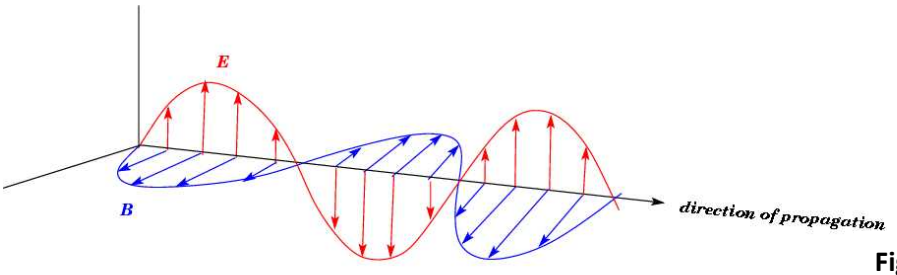


Figure 11.19

The existence of such waves was predicted by James Clerk Maxwell, who in the 1860s synthesized many of the concepts in Chapter 9 and developed the theory of electromagnetism. Electromagnetic waves are produced by oscillating electric charges, in particular electrons.

Because the E-field and B-field in an electromagnetic wave are oscillating at right angles to the direction of propagation, they are **transverse waves** of the type discussed in Section 5.8 and they share all the basic properties of transverse waves. **The direction of the E-field vector is termed the direction, or axis, of polarization.**

In extended light sources, such as an incandescent bulb, each section of the source acts independently of the others; the E-fields of the light emitted by the different sections are therefore vibrating in different directions randomly. In that case we say that the light is **unpolarized**. Other sources, for example a laser, emit light such that the E-field is vibrating in a single direction; such light is **polarized**. When light is unpolarized, it can be viewed as consisting of equal parts of light oscillating in any two mutually perpendicular directions, for example the x- and y-directions (Figure 11.20a). This is simply because all the E-field vectors can be resolved into x- and y-components. In contrast, when light is polarized it oscillates in one direction only (Figure 11.20b).

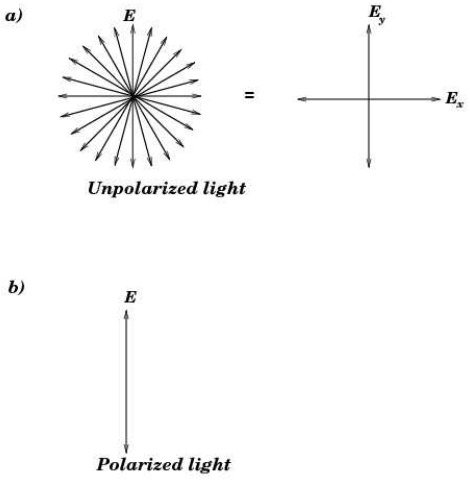


Figure 11.20

The molecules in a crystal of Iceland spar are also aligned in one direction. **Only light whose polarization axis is correctly aligned with the crystal axis can pass through it.** Any material that behaves like Iceland spar is termed a **polarizer**. Sunglasses are the most common example.

Light reflected off a window tends to be polarized. Thus, its E-field vector, **E,** oscillates in only one direction. If the light passes through a polarizer, **E** will make some angle θ with the polarizer axis (Figure 11.21).

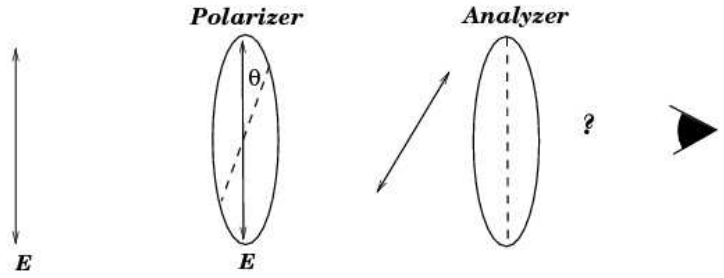


Figure 11.21

Only the component of **E** lying along the polarizer axis can be transmitted. The figure shows that the amplitude of the transmitted E-field is therefore $E\cos\theta$. Now, light intensity I is defined as power per unit area. We state without proof that the intensity of light is proportional to the square of the amplitude, implying that $I \propto \cos^2\theta$, or $I = I_o\cos^2\theta$, where I_o is the maximum intensity. We recognize this expression as Malus' law, Eq. (11.9). If the light is then passed through a second polarizer (usually called the **analyzer**) whose axis is crossed at right angles with respect to the first polarizer, then no light gets through, as Christiaan Huygens observed.

Example 11.8: Light bulbs. Light from a light bulb passes through a polarizer, then through an analyzer whose axis is rotated 15° with respect to the polarizer axis. What fraction of the original intensity gets through?

A light bulb emits unpolarized light, as in Figure 11.20a. The light is vibrating in all directions randomly and can be viewed as consisting of equal parts light vibrating along the polarizer axis and perpendicular to it. Therefore, only half the original intensity I_o is transmitted by the polarizer. (This is the purpose of polaroid sunglasses, to cut down the intensity by half.) Because the analyzer is rotated by 15°, Malus's law implies that the final transmitted intensity is $I = (1/2)I_o\cos^2 15 = 0.466I_o$.

Further aspects of polarization will be explored in the problems.

11. Interference and Diffraction of Light

We have discussed interference of waves in Section 5.13. Like other waves, light undergoes interference, although because the wavelength of light is so short, historically this was not obvious. Indeed, because interference is the defining property of waves—but not particles—until it could be demonstrated that light underwent interference, it was not obvious that light was a wave at all. The experiment that demonstrated the interference of light is universally known as Young's double-slit experiment, or **Young's two-slit experiment**, although the evidence that Thomas Young actually performed it is inconclusive. Whoever first did it, it is considered perhaps the most important experiment in the history of science. Young published a description of a hypothetical experiment in 1807.

We encountered the two-slit experiment in Problem 5.19:

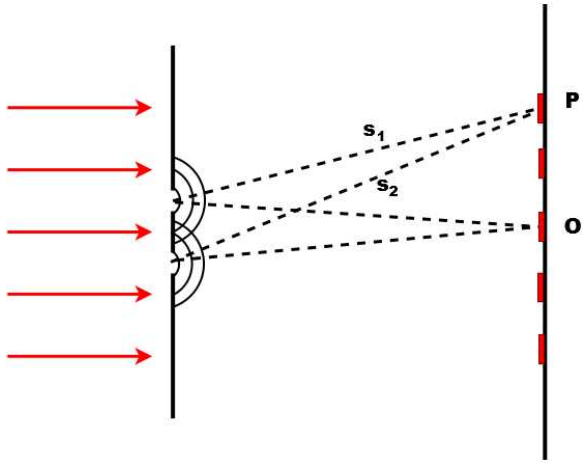


Figure 11.22

A collimated beam of light of a single frequency (preferably from a laser) is passed through two narrow slits in an opaque screen (Figure 11.22). A laser beam displays a high degree of **coherence**: the crests and troughs of the waves line up in a fixed pattern that does not change with time. On a distant wall, one then observes an **interference pattern**: a series of alternating bright and dark stripes, or **fringes** (thick bars = bright). *It is only with coherent waves that interference takes place.* If the waves are of different frequencies and the crests and troughs arrive at the slits at random times, there will be no constant superposition of waves and no interference pattern will be observed. Lasers are the gold standard of coherent light. Incandescent bulbs produce incoherent light.

According to **Huygens' principle** (Problem 5.19), when illuminated by a coherent light beam, the slits *become new sources of light with the same phase as the incoming beam*. In other words, the two slits themselves emit coherent light. If the collimated beam illuminates the slits head on, the waves at the slits will be in phase. At a point O on the far wall midway between the slits, the light from each slit has traveled the same distance. The waves remain in phase at that point, undergo constructive interference and a bright fringe appears there.

At other points P the light from the slits will have traveled different distances s_1 and s_2 and therefore the path difference will be $\Delta s = s_2 - s_1 \neq 0$.

Tip: The path difference is *always* the crucial quantity in interference problems.

From Chapter 5 we know that if Δs is an integral number of wavelengths, the interference at P will be constructive and a bright spot will appear there. Therefore, if λ is the wavelength of light, the condition for constructive interference can be written

$$\Delta s = m\lambda \qquad m = 0, 1, 2 \ldots$$

There are several equivalent methods to determine the path difference Δs. In Figure 11.23, L is the distance from the screen to the wall, which may be several meters. The separation between the two slits is $2a$, which is usually a fraction of a millimeter. The distance from the midpoint on the wall to any given bright spot is marked y, which is usually several centimeters.

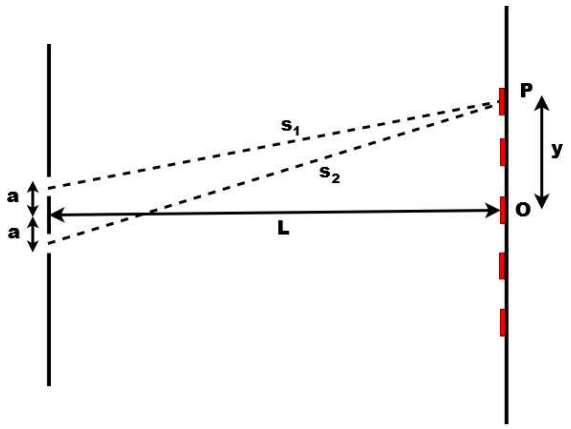


Figure 11.23

From the Pythagorean theorem,

$$s_1^2 = L^2 + (y - a)^2$$

and

$$s_2^2 = L^2 + (y + a)^2.$$

Squaring out the parentheses and subtracting the first equation from the second gives

$$s_2^2 - s_1^2 = 4ay.$$

The left side can be factored to give

$$(s_2 + s_1)(s_2 - s_1) = 4ay\,.$$

However, remember that the horizontal distance L (meters) is much larger than the vertical distance y (centimeters). Thus, s_1 and s_2 are each very nearly L itself and $(s_1 + s_2) \approx 2L$. Consequently,

$$\Delta s = s_2 - s_1 \approx \frac{2ay}{L}.$$

But $2a$ is the separation between the slits, which we can call d, and the condition for constructive interference then becomes in this approximation

$$\Delta s = \frac{dy}{L} = m\lambda \quad ;\ m = 0,1,2\ldots \qquad (11.10)$$

An alternative way of seeing this is to assume that the wall is so far away that s_1 and s_2 are essentially parallel lines (Figure 11.24). Then the small triangle indicated in the figure is a right triangle and the path difference Δs is the extra distance light traveling along s_2 must travel. The figure shows $\Delta s = 2a \sin\theta$. (**Exercise:** Convince yourself of this.) However, because the wall is far away, $y \ll L$. Then one can use the small angle approximation[2] $\sin\theta \approx \theta \approx y/L$. Thus, $\Delta s \approx dy/L$, as before.

[2] For angles $\theta \ll 1$ rad, $\sin\theta \approx \tan\theta \approx \theta$ (in radians) and $\cos\theta \approx 1$.

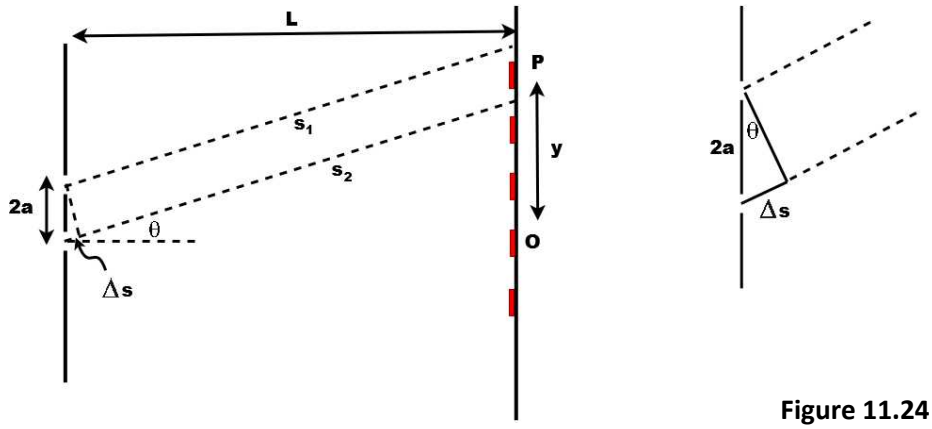


Figure 11.24

For destructive interference, recall that the condition is that the two waves must be an odd integer of half wavelengths out of phase. Eq. (11.10) then becomes

$$\frac{dy}{L} = \left(m + \frac{1}{2}\right)\lambda \quad ; \; m = 0, 1, 2 \ldots \qquad (11.11)$$

where y is now the distance to a dark spot.

Example 11.9: Laser pointers. Light from a commercially available laser pointer is shone through two slits separated by 0.119 mm and an interference pattern is observed on a screen 4 m away. The distance on the screen to the second bright spot from the center is 3.5 cm. What type of laser is it?

Using Eq. (11.10) with $m = 2$, we have $\lambda = dy/2L = (1.2 \times 10^{-4})(3.5 \times 10^{-2})/8 = 525$ nm. This is in the green part of the spectrum and is close to the wavelength (532 nm) advertised for Neodymium-doped Yttrium lithium fluoride laser pointers, so it is probably that sort of laser.

As stated in Section 10, light intensity is defined as power per unit area. If one measures the intensity of the bright spots on the wall due to two-slits, one finds a characteristic pattern called the **two-slit interference pattern,** or **two-slit diffraction pattern**, drawn in Figure 11.25.

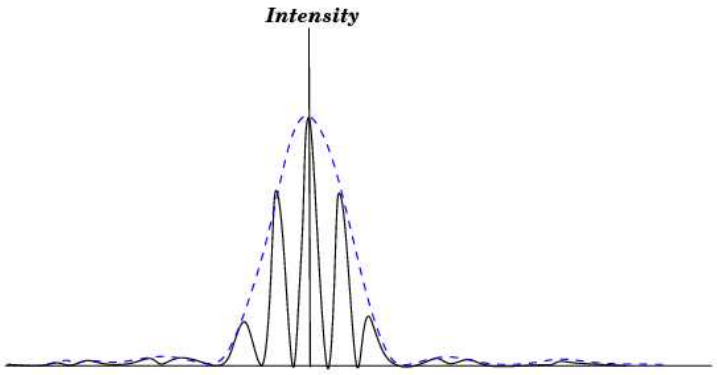


Figure 11.25

The intensity is a maximum at the central bright spot and falls off in the complicated but easily recognizable pattern shown. As an **Exercise**, you should mark the position of the maxima and minima on the x-axis as required by Eqs. (11.10-11.11). Do this either in terms of y/L or θ (see below).

The dashed envelope drawn in the figure turns out to be the **diffraction pattern** due to a single slit. The term **diffraction** refers to the bending of light around the edges of objects. As the word has come to be used, most physics texts refer to two-slit interference but single-slit diffraction. However, in both cases light is being bent around the edges of the slits, which results in the light interfering with itself. Many physicists employ the terms interchangeably and refer to "two-slit diffraction." In any case, light passing through a single slit is bent, or diffracted, and produces an interference pattern on a distant screen that displays maxima and minima. We state without proof that for a slit of width a, the *minima* of the diffraction pattern occur when

$$a \sin\theta = m\,\lambda \quad m = 0,1,2 \ldots \quad (11.12)$$

The angle θ is the angle between the central peak and the given minimum, as shown in Figure 11.26.

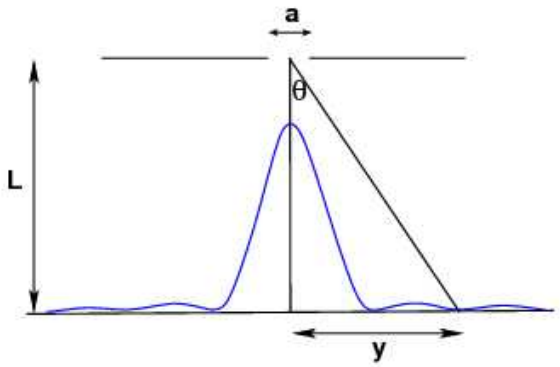


Figure 11.26

If the screen is far enough away from the slit, then $\sin\theta \approx y/L$, as above, and the condition for the minima becomes $ay/L = m\lambda$. *This expression is of exactly the same form as Eq. (11.10) except that there d was the separation between the two slits and here a is the width of the single slit, and here the formula gives the condition for the intensity minima*. **Warning: Do not confuse the two formulas.**

The reason that the envelope of a single-slit pattern appears in the double-slit pattern in Figure 11.25 is far from obvious. Ultimately, the behavior is due to the fact that the slits are not infinitely narrow but have a finite width. If the slits were infinitely narrow, the light they produced would be of uniform intensity and would add together to produce interference maxima of uniform intensity as well. The two-slit pattern in Figure 11.25 would then consist of peaks of equal height. However, the slits have a finite width and thus each one is producing its own single-slit pattern, which has the varying intensity of Figure 11.26. In places where the single-slit pattern becomes zero, the double-slit pattern must also. So you get a combination pattern. The situation is similar to that in Example 5.12, where we produced a standing wave by adding two traveling waves together, which turned out to be equivalent to multiplying two oscillating functions together. You can say that the single-slit pattern multiplies the double-slit pattern, or if you want to impress your friends, you can say that the single-slit pattern **modulates** the double-slit pattern.

Example 11.10: Missing maxima. A two-slit interference pattern is missing its third maximum from the central bright spot. What is the ratio of the slit width to the slit separation?

Assume that the third maximum of the two-slit pattern falls under the first minimum of the one-slit pattern, as in Figure 11.25. Then from Eq. (11.10) we have $dy/L = 3\lambda$, but from Eq. (11.12) we have $ay/L = 1 \cdot \lambda$. Dividing the second equation by the first gives $a/d = 1/3$, so the slits are separated by three times their width.

12. Thin Film Interference

At some point in your life you have undoubtedly noticed the rainbow of colors on a soap bubble, on an oil slick or on a camera lens. These phenomena are also the result of the interference of light waves. Precisely, they are due to the interference of light that is reflected off both the front and back surfaces of **thin films**—films whose thickness is approximately that of a wavelength of light. The use of thin films has become indispensable in creating optical systems that transmit or reflect only desired wavelengths, both in the visible and nonvisible regions of the electromagnetic spectrum. The AP exams frequently contain questions regarding thin-film interference. All such problems are similar to the following example, which serves to explain the concept.

Consider a film of motor oil floating on water. The index of refraction of the motor oil is $n = 1.50$ and the thickness of the film is $d = 400$ nm. If white light is normally incident on the oil, what color will it appear?

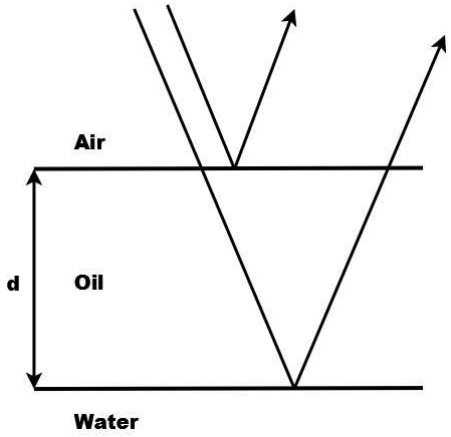


Figure 11.27

The situation is illustrated in Figure 11.27, where the angles have been exaggerated; we are to assume that the rays are striking the oil film nearly head on. The color of any object is determined by the wavelength of light that it *reflects*. In this case, some of the light is reflected off the top surface, while some is transmitted through the oil to bounce off the bottom surface. A ray reflected off the bottom surface will return to the oil-air interface and interfere with a ray that merely bounced off the top. If the interference is destructive at a certain wavelength, the color corresponding to that wavelength won't be seen. Only wavelengths that undergo constructive interference are visible to an observer looking at the film.

The path difference between the top and bottom ray is $2d$ for normal incidence. You might reasonably conclude that if $2d$ is an integral number of wavelengths, constructive interference for that wavelength will take place. Unfortunately, two phenomena complicate the situation. First, the refractive index for oil is greater than 1. According to Eq. (11.4) the wavelength of light in the oil is then shortened from its vacuum value λ_o to $\lambda_n = \lambda_o/n$. Since the light reflected off the bottom surface is traveling in the oil, the appropriate wavelength to use is λ_n, not λ_o. (When the light enters the air again, the wavelength expands to λ_o once more, but whether a crest or trough appears at the oil-air interface, depends on how far it has traveled in the oil.)

Revised Tip: In any optics problem, the optical path length, or number of wavelengths in a medium, is the important quantity, not the physical path length.

(The two-slit experiment took place in air; the optical path length and the physical path length were the same.)

You might now reasonably conclude that setting $2d = m\lambda_n$ is the correct condition for an interference maximum, but this also turns out not to be correct. We state without proof that

When a wave traveling in a medium with refractive index n_1 is reflected from a medium with a refractive index $n_2 > n_1$, it undergoes a phase change of π radians, or 180°.

Remember that the crests and troughs of a sinusoidal wave (Figures 5.7 and 5.11) are separated by π radians. Thus "a phase change of π radians" means that the crests and troughs have exchanged positions. In the current situation, the light is initially traveling in air with $n_1 = 1$. The ray reflected off an oil film with $n_2 > 1$ must undergo a phase flip. The ray reflected off the bottom surface does *not* undergo a phase flip because $n_{H_2O} < n_{oil}$.

Consequently, in this example the correct condition for constructive interference is

$$2d = \left(m + \frac{1}{2}\right)\lambda_n .$$

We are given that $d = 400$ nm and $n = 1.5$. Thus, the wavelengths $\lambda_o = n\lambda_n$ that will be most visible are

$$\lambda_o = \frac{1200}{m + \frac{1}{2}} \text{ nm.}$$

For $m = 0$, we find $\lambda_o = 2400$ nm, which is well in the infrared part of the spectrum and not visible to the human eye (Figure 11.28). For $m = 1$ we get $\lambda_o = 800$ nm, still in the infrared. For $m = 2$, we find $\lambda_o = 480$ nm, in the blue part of the spectrum. For $m > 2$ the results are less than 400 nm. Hence, they are in the ultraviolet part of the spectrum and also invisible. Thus, we expect that such an oil slick should appear mostly blue or purple.

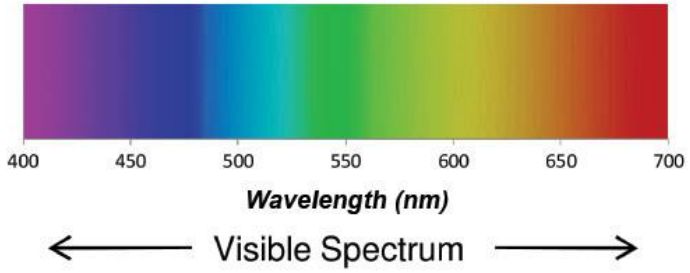


Figure 11.28

The take-home message from this example is that **when thin-films are involved the condition for constructive or destructive interference varies from problem to problem**. You must examine each situation carefully to determine the appropriate interference conditions. For example, suppose the oil film were spread out on a piece of glass with $n = 1.65$. In that case there would be a phase flip at both the top and bottom surfaces and the condition for constructive interference would be $2d = m\lambda_n$. With the same numbers as above, we find that the visible wavelengths are given by $m = 2$ and $m = 3$, resulting in $\lambda_o = 600$ nm and $\lambda_o = 400$, the latter being on the lower end of the visible spectrum. Therefore, we expect the oil slick to appear mostly greenish-yellow.

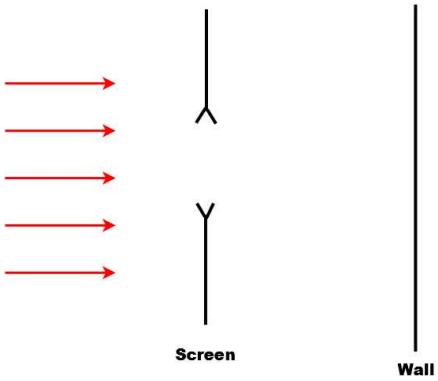
On the other hand, supposed you wanted to design a lens coating that transmitted green light. The complementary color of green is red and so the strategy is to design a coating that reflects red light, say 630 nm. If we assume that the index of refraction of the coating is greater than 1 but less than the glass on which the coating is applied, there are again two phase flips and the condition that red is reflected is the same as just used: $2d = m\lambda_n = m\lambda_o/n$. When n is given, this equation tells you the allowed thickness for the coating that will reflect light of 630 nm. For $n = 1.4$, the thinnest such coating is given by $m = 1$, or $d = 225$ nm.

Further variations on this theme will be found in the problems.

OPTICS PROBLEMS

Problems 1-20 are at the subject-test level. The remaining problems are at the AP level.

Questions 1 - 3 refer to the following figure:



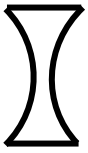
1. A beam of parallel light rays is incident on an opaque screen that has a large opening in it. The opening is fitted with clamps designed to hold optical devices. If you want to produce a tiny spot of light on a distant wall, you should place in the clamps which of the following devices?

(A) A prism

(B) A narrow slit in a piece of cardboard

(C) A rectangular piece of glass

(D) A converging lens

(E) A diverging lens 

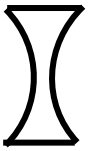
2. A beam of parallel light rays is incident on the same screen in the previous problem. If you want to produce a diffraction pattern on the distant wall, you should place in the clamps which of the following optical devices?

(A) A prism

(B) A narrow slit in a piece of cardboard

(C) A rectangular piece of glass

(D) A converging lens

(E) A diverging lens 

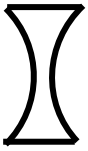
3. A beam of white light is incident on the same screen as in problems 1-2. If you want to produce a spectrum on a distant wall, you should place in the clamps which of the following?

(A) A prism

(B) A narrow slit in a piece of cardboard

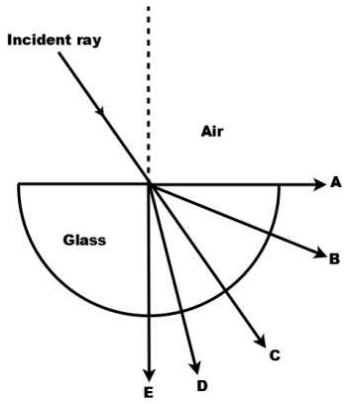
(C) A rectangular piece of glass

(D) A converging lens

(E) A diverging lens 

4. A ray of light passing through air is incident on a semicircular piece of glass, as shown below.

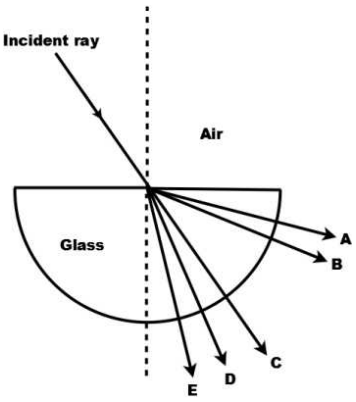
Which of the drawn rays might represent the path of the light after it enters the glass?



(A) A
(B) B
(C) C
(D) D
(E) E

5. A ray of white light passing through air is incident on a semicircular piece of glass, as shown below.

Which of the drawn rays might represent the paths of red and blue light after the ray enters the glass?



(A) A for red; B for blue
(B) B for red; A for blue
(C) C for both red and blue
(D) D for blue; E for red
(E) E for blue; D for red

6. A beam of light traveling through a piece of glass emerges into air. Which of the following statements is true as the light exits the glass?

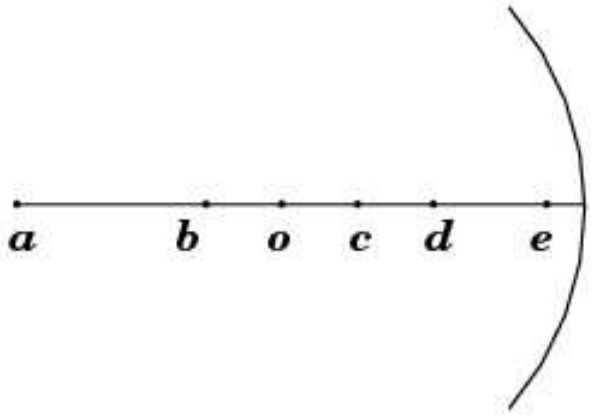
i. The speed of light stays the same; ii. The light speeds up; iii. The light slows down;
iv. The frequency of the light stays the same; v. The frequency of the light increases; vi. The frequency of the light decreases; vii. The wavelength of the light stays the same; viii. The wavelength of the light increases; ix. The wavelength of the light decreases.

(A) i, iv, vii
(B) i, v and vii
(C) ii, iv and viii
(D) ii, v, ix
(E) iii, iv and ix

7. A concave spherical mirror has a center of curvature marked *o*, as shown below.

Which of the other indicated points could be the mirror focus?

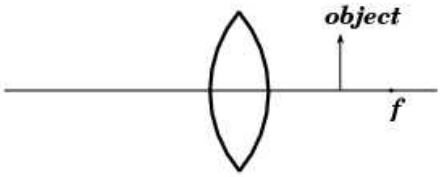
(A) a
(B) b
(C) c
(D) d
(E) e



8. An object is placed inside the focus of a convex lens, as shown below.

The image produced is

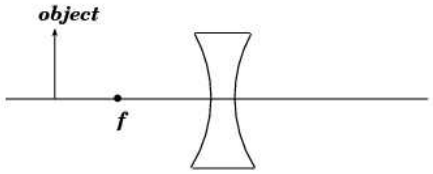
(A) virtual and upright.
(B) real and upright.
(C) virtual and inverted.
(D) real and inverted.
(E) no image is formed.



9. An object is placed outside the focus of a diverging lens, as shown below.

The image produced is

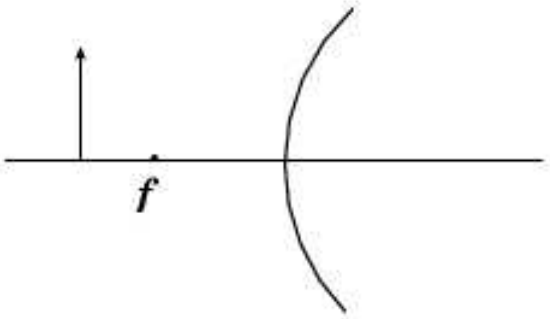
(A) virtual, upright, on the object side of the lens
(B) virtual, upright, on the opposite side of the lens from the object
(C) virtual, inverted, on the opposite side of the lens from the object
(D) real, inverted, on the object side of the lens
(E) no image is formed



10. An object is placed outside the focus of a convex mirror, as shown below.

The image produced is

(A) virtual, upright, in front of the mirror
(B) virtual, upright, behind the mirror
(C) virtual, inverted, behind the mirror
(D) real, inverted, in front of the mirror
(E) no image is formed

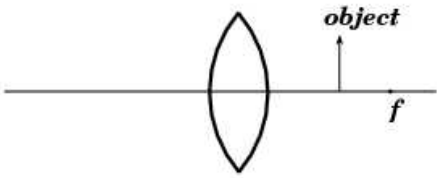


11. At what distance d_o should an object be placed in front of a converging lens such that the image is at the same distance on the other side of the lens?

    (A) $d_o = f$
    (B) $d_o = \infty$
    (C) $d_o = f/2$
    (D) $d_o = 2f$
    (E) It cannot be done

12. An object is within the focal length of a convex lens, as shown below.

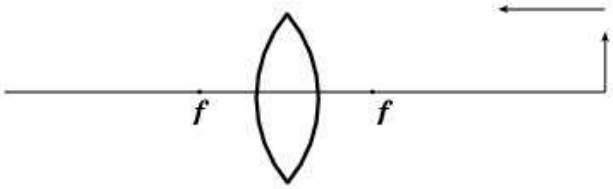
    As the object is moved from the focus towards the lens, which of the following statements is true?



    (A) The image increases in size and moves closer to the lens.
    (B) The image decreases in size and moves closer to the lens.
    (C) The image increases in size and moves farther from the lens.
    (D) The image decreases in size and moves farther from the lens.
    (E) The answer depends on the focal length of the lens.

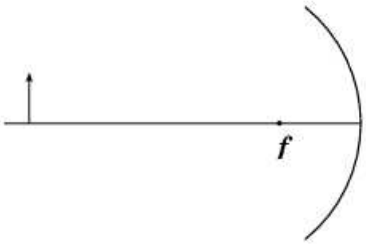
13. An object initially at infinity is brought closer to a converging lens.

    As the object is moved,



    (A) the inverted image moves closer to the far focus and becomes larger.
    (B) the inverted image moves farther out from the far focus and becomes larger.
    (C) the upright image moves closer to the far focus and becomes larger.
    (D) the upright image moves farther out from the far focus and becomes larger.
    (E) the inverted image moves farther out from the far focus and becomes smaller.

Questions 14 - 16 refer to the following figure:



14. An object initially at infinity is brought closer to the focus of a concave mirror. As the object moves in,

    (A) the inverted image moves closer to the focus and becomes larger.
    (B) the inverted image moves farther out from the focus and becomes larger.
    (C) the upright image moves closer to the focus and becomes larger.
    (D) the upright image gets farther out from the focus and becomes larger.
    (E) the inverted image gets farther from the focus and becomes smaller.

15. As in the previous problem, an object initially at infinity is brought through the focus of a concave mirror. As the object is moved,

    (A) a real, inverted image becomes an upright, virtual image on the object side of the mirror.
    (B) a real, inverted image becomes an upright, real image on the object side of the mirror.
    (C) an upright, virtual image on the object side of the mirror becomes an upright, virtual image behind the mirror.
    (D) a virtual, inverted image on the object side of the mirror becomes an upright, virtual image behind the mirror.
    (E) a real, inverted image becomes an upright, virtual image and appears behind the mirror.

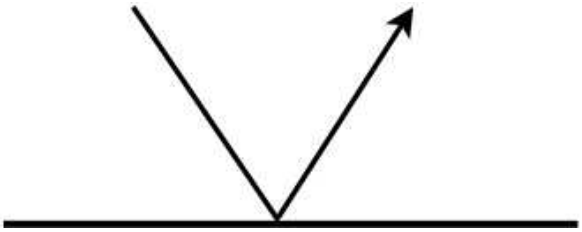
16. An object initially at the focus of a concave mirror is brought closer to the mirror. As the object is moved toward the mirror's surface,

    (A) the image gets larger until it is of infinite size when the object reaches the mirror surface.
    (B) the image gets smaller until it is twice the object size when the object reaches the mirror surface.
    (C) the image gets smaller until it becomes equal to the object size when the object reaches the mirror surface.
    (D) the image gets larger until it is twice the object size when the object reaches the mirror surface.
    (E) the image stays the same size.

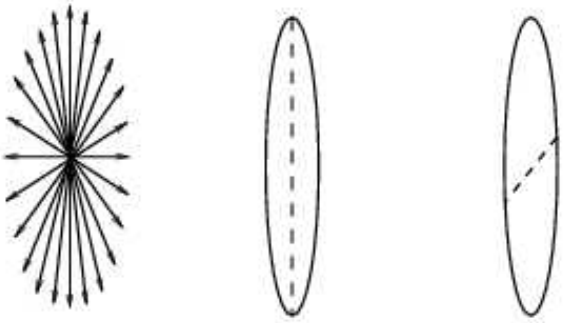
17. A beam of light is reflected off a pane of glass, as shown below.

    Which statement is true?

    (A) An inverted image will be formed at infinity.
    (B) The light will be partially polarized.
    (C) The speed of light will be increased.
    (D) The intensity of light will be increased.
    (E) The light will be focused.



Questions 18 - 20 refer to the following figure:



18. A beam of unpolarized light is passed through two polarizers. If the polarization axis of the second polarizer is at an angle of 90° with respect to the axis of the first polarizer, then the intensity of light seen by someone located between them is

    (A) the intensity of the original light.
    (B) one-half the intensity of the original light.
    (C) one-quarter the intensity of the original light.
    (D) one-eighth the intensity of the original light.
    (E) zero.

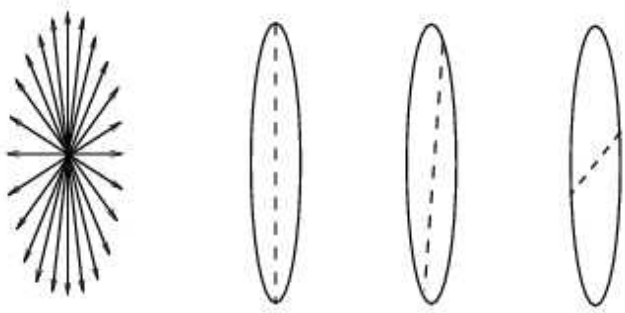
19. A beam of unpolarized light is passed through two polarizers. If the polarization axis of the second polarizer is crossed at an angle of $90°$ with respect to the axis of the first polarizer, then the intensity of light seen by someone located to the right of the second polarizer is

    (A) the intensity of the original light.
    (B) one-half the intensity of the original light.
    (C) one-quarter the intensity of the original light.
    (D) one-eighth the intensity of the original light.
    (E) zero.

20. A beam of unpolarized light is passed through two polarizers. If the polarization axis of the second polarizer is at an angle of $45°$ with respect to the axis of the first polarizer, then the intensity of light seen by someone located to the right of the second polarizer is

    (A) the intensity of the original light.
    (B) one-half the intensity of the original light.
    (C) one-quarter the intensity of the original light.
    (D) one-eighth the intensity of the original light.
    (E) zero.

21. A beam of unpolarized light is passed through two polarizers whose axes are crossed at $90°$ with respect to one another. A third polarizer is inserted between the two whose axis is aligned at $45°$ with respect to both, as shown below.



    The intensity of light seen by someone located to the right of all three polarizers is

    (A) the intensity of the original light.
    (B) one-half the intensity of the original light.
    (C) one-quarter the intensity of the original light.
    (D) one-eighth the intensity of the original light.
    (E) zero.

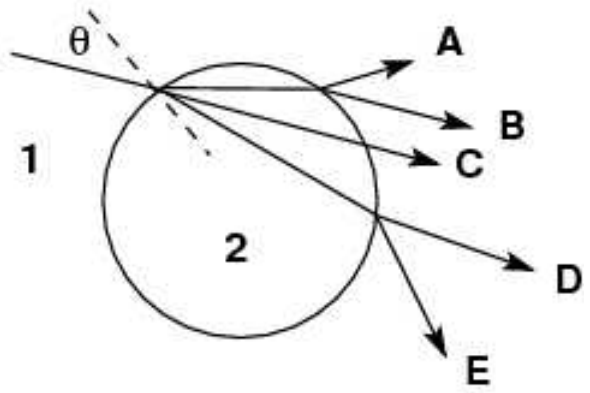
22. A beam of light in water ($n_1 = 4/3$) strikes an interface with a piece of glass ($n_2 = 1.5$). The critical angle at which total internal reflection takes place is

    (A) $0°$
    (B) $48.6°$
    (C) $62.7°$
    (D) $90°$
    (E) Total internal reflection cannot take place

23. A beam of light in glass ($n_1 = 1.5$) strikes an interface with water ($n_2 = 4/3$). The critical angle at which total internal reflection takes place is most nearly

(A) 0°
(B) 48.6°
(C) 62.7°
(D) 90°
(E) Total internal reflection cannot take place

Questions 24 - 25 refer to the following figure:



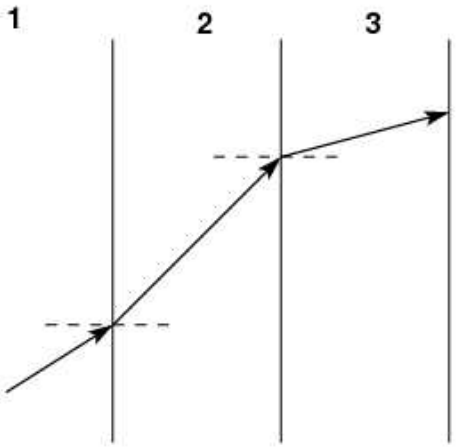
24. A beam of light originating in medium 1, strikes a sphere of medium 2 at an angle θ from the normal of the sphere, as above. If medium 1 is air and medium 2 is glass, which is a possible trajectory for the light beam?

(A) A
(B) B
(C) C
(D) D
(E) E

25. If in the previous problem, medium 1 is glass and medium 2 is air, which is a possible trajectory for the light beam?

(A) A
(B) B
(C) C
(D) D
(E) E

26. A beam of light passes from a medium 1 with a refractive index n_1 to a medium 2 with refractive index n_2 to a medium 3 with refractive index n_3, as shown on the right.



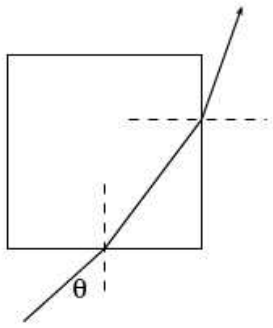
Which ranking of the refractive indexes is correct?

(A) $n_2 < n_1$; $n_2 = n_3$
(B) $n_1 < n_2 < n_3$
(C) $n_3 < n_2 < n_1$
(D) $n_2 < n_1 < n_3$
(E) $n_2 < n_1$; $n_1 = n_3$

27. A beam of light in air ($n = 1$) strikes the horizontal face of a glass cube ($n > 1$) at an angle $0 < \theta < 90°$, as shown. The beam proceeds to the vertical face of the cube.

Which statement is true?

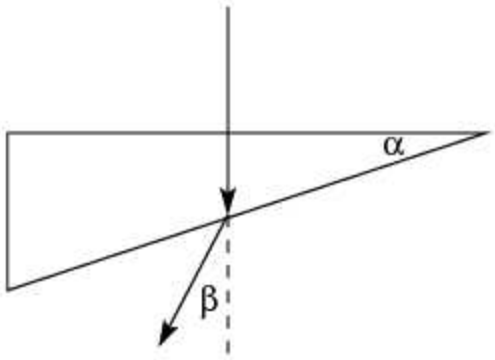
(A) The beam will never exit the cube.
(B) The beam will always exit the cube.
(C) The beam will exit the cube if $n < \sqrt{2}$.
(D) The beam will exit the cube if $\theta > 45°$.
(E) The beam will exit the cube if $\theta < 45°$.



28. A beam of light in air ($n = 1$) is normally incident on the horizontal face of a triangular prism, as shown below. The emerging light makes an angle $\beta = 40°$ from the light's original direction.

If the prism angle is $\alpha = 20°$, the prism's refractive index n is most nearly

(A) 1.5
(B) 2.53
(C) 0.75
(D) 1.4
(E) 2.10



29. A beam of light in air ($n = 1$) is normally incident on a prism of the same shape as in the previous problem. However, this time it is incident on the left, vertical face. The index of refraction of the prism is $n = 2.25$.

The angle at which it emerges from the prism relative to its original direction is

(A) It does not emerge; there is total internal reflection.
(B) 26.4°
(C) 63.6°
(D) 0°
(E) 20°

Free-response problems

30. You are given a thin converging lens with a focal length of 10 cm.

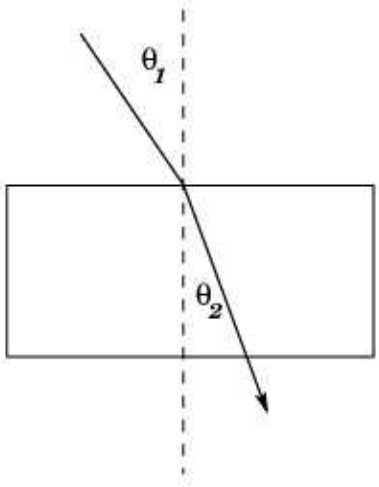
    a) An object is placed at 6 cm from the lens. Draw a ray diagram for the situation.

    b) Calculate the image distance. Does the calculated distance agree with the distance given by the ray diagram? Is the image real or virtual?

    c) The object is moved out to 8 cm. Draw a ray diagram for the situation.

    d) Calculate the image distance. Does the calculated distance agree with the distance given by the ray diagram? Is the image real or virtual. Is it larger or smaller than the image from part (a)?

31. You are attempting to determine the index of refraction of a block of an unknown material. You shine a light beam at various angles θ_1 from the normal to the block and measure the refraction angle θ_2, as shown.



You collect the following data:

θ_1 (deg)	θ_2 (deg)
10	5.0
20	12.5
30	18.0
40	24.0
50	28.6
60	33.0
70	36.0
80	40.0

    a) Determine the index of refraction of the material by two methods.

    b) Which points would you say represent bad measurements?

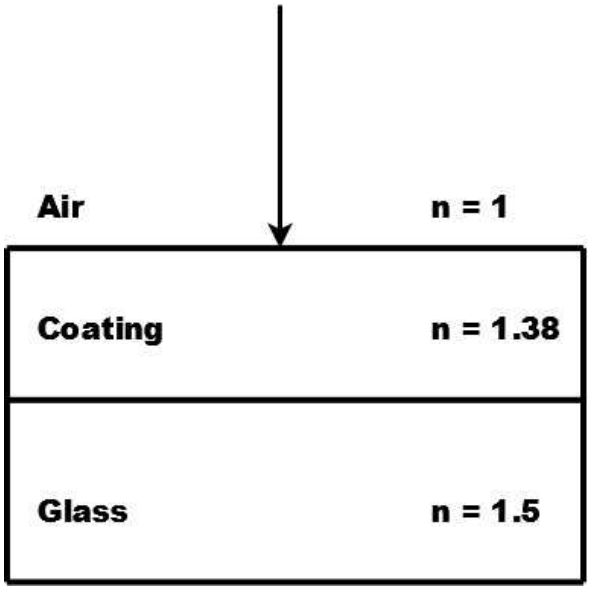
32. You stand at a distance x from a plane (flat) wall mirror.

    a) What is the focal length of the mirror?

    b) How far behind the mirror is your image formed? Is the image real or virtual?

    c) How tall must the mirror be in order that you see your full body?

33. Magnesium fluoride, MgF_2, is a transparent substance widely used for optical coatings, especially in space telescopes. White light is normally incident on a glass lens coated with a thin film of magnesium fluoride, as shown below. The refractive index of MgF_2 is $n = 1.38$ and the refractive index of the glass is $n = 1.5$.



Light reflected directly upward from the coating is observed to be predominantly green ($\lambda = 525$ nm).

a) What is the frequency of the green light in air? In the coating?

b) What is the wavelength of the green light in the coating?

c) What is the minimum thickness of the coating that allows the green light to be reflected?

d) If instead of reflecting green light, you wish to design a coating to *eliminate* reflections at 525 nm, what should be the minimum thickness of the coating?

e) Suppose you wished the lens to transmit red light at 670 nm and reflect blue light at 500 nm, what is the minimum thickness of the coating?

f) Suppose in parts (a)-(c) that the light source is a bulb and that you gradually move to the left or right as you observe the reflected light. You see that the reflected light changes away from green. How does the color change and what is the explanation for the phenomenon?

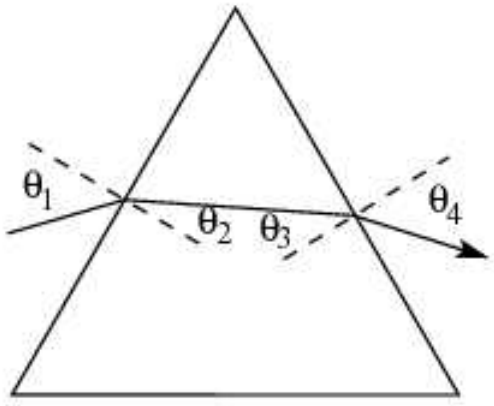
34. A prism in the form of an equilateral triangle (below) has a refractive index $n = 1.6$.

A beam of light is incident from air onto the prism at $\theta_1 = 50°$.

a) What is θ_2?

b) What is θ_4?

c) How can the setup be changed so that total internal reflection takes place?

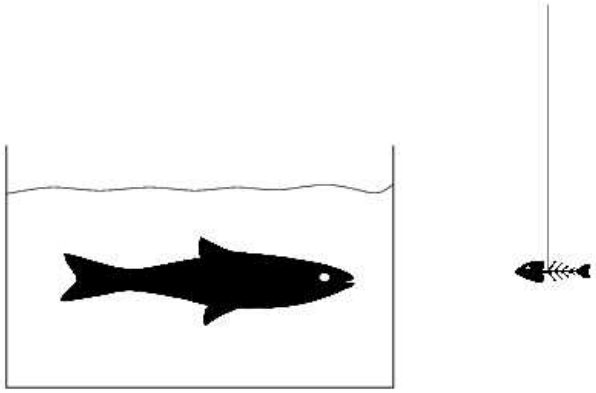


d) A thin film of magnesium fluoride ($n = 1.38$) is painted on the left side of the prism to reduce reflection of red light ($\lambda = 630$ nm) at normal incidence. What is the wavelength of the red light inside the film?

e) What should be its thickness to minimize the reflection of red light?

f) If the light is glancing off the prism at an angle, rather than at normal incidence, do you expect the reduction in the reflection of light to be more effective at longer or shorter wavelengths, or to remain the same?

35. A big fish in an aquarium full of water greedily eyes a little fish hanging from a mobile outside the tank in air. The index of refraction of the water is $n = 4/3$. Ignore any effect of the glass wall of the aquarium.

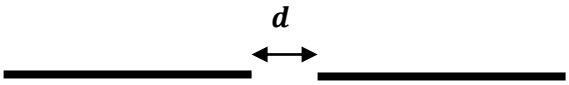


a) Assuming that the big fish's eye is next to the aquarium wall, does the little fish appear nearer, farther or at its actual distance? Justify your answer.

b) At what distance from the aquarium wall does the little fish appear? (You may assume small angles.)

c) Is the image that the big fish sees of the little fish larger, smaller or of the same size as the actual fish? Justify your answer.

36. You are given an opaque sheet of metal with a narrow slit in it, of width d (below).



You are also given the following:

1) A digital timer; 2) A meter stick; 3) A metric ruler; 4) A light bulb with power cord

5) A laser pointer ($\lambda = 630$ nm); 6) A prism; 7) A projection screen; 8) A mirror

Your mission is to measure d with the proper pieces of the above equipment.

a) Check off the equipment you would use.

b) Draw a diagram of the experimental setup.

c) Describe how you would perform the experiment.

d) What would be the difference in the experiment if the metal had two narrow slits cut into it and you were asked to determine the distance between the slits?

37. A sliver of mica ($n = 1.58$) of thickness h is placed over one slit in a two-slit arrangement, as shown below.

What had previously been the central bright spot on the distant viewing screen is now observed to be occupied by the sixth bright fringe. If the wavelength of light is $\lambda = 550$ nm, what is the thickness of the mica?

12. MODERN PHYSICS REVIEW

The material covered by this chapter is intrinsically at the AP level. However, the subject test expects a "scientific literacy" of advanced topics. Sections 9.1-9.3 and Section 9.10 are at the subject-test level. Sections 9.4-9.9 are at the AP level. The subject-test taker is advised to read at least the introductory paragraphs on the AP sections.

1. Essential Terms

In the early twentieth century, two developments revolutionized physics. One was the creation of **special relativity** (Section 6.4) in 1905 by Albert Einstein and the other was the invention of **quantum mechanics** by Max Planck. Quantum mechanics was born on December 14, 1900 when Planck postulated that energy could not come in arbitrary amounts, as we have assumed throughout this book, but was **quantized**, or appeared in discrete units.

One of the first demonstrations of quantization was the **photoelectric effect**, in which **electrons** are ejected from a piece of metal when it is struck by light. The photoelectric effect was explained by Einstein in 1905 by assuming that the light came in discrete energy packets known as **quanta**. Later, light quanta, or particles, became known as **photons**. The photoelectric effect and other experiments therefore established that light, which had been thought to be a wave, also had particle properties. In 1924 Louis de Broglie hypothesized that matter itself had wave properties—specifically a wavelength and a frequency. The proposal was soon verified by experiment. The fact that waves can behave like particles and particles can behave like waves is often referred to as the **wave-particle duality**.

The quantum hypothesis, that energy, momentum and angular momentum can attain only certain values, underlies all **atomic physics**. At an elementary level, atomic physics largely concerns the production of **atomic spectra**—the emission of light as electrons change from one allowed **energy level** to another. The analysis of atomic spectra proved the key to understanding the structure and composition of stars.

Quantum mechanics also governs nuclear physics. The atomic **nucleus** consists of **protons** and **neutrons**. A given chemical **element** is determined by the number of protons in the nucleus. However, an element may appear in various **isotopes**, which differ from each other in the number of neutrons the nucleus contains. At the level of elementary courses, one of the most important results of nuclear physics is the phenomenon of **radioactivity**, which is the emission of particles by the nucleus. During the process, the nucleus loses energy by $E = mc^2$ and decays into a different isotope of the original element, or into a different element altogether. Atomic and nuclear physics are together subsumed under the more general topic of **particle physics**, which deals with all **subatomic particles**, not only the electron, proton and neutron, but the photon, **neutrino** and many others.

Since 1900, physicists have had to distinguish between **classical physics** (no quantum mechanics involved; sometimes no relativity) and quantum physics. While it is true that many aspects of quantum mechanics are extremely weird, leaving physicists scratching their heads, the theory has no known violations and in our times has become an essential tool in understanding nature. It also underlies the entire modern electronics industry and according to some estimates about one-third of the US economy is now based on quantum mechanics.

The mathematics of full quantum theory is far above the level of this text. However, the SAT subject test does require a qualitative understanding of all the above topics. The AP exam generally includes at least one question on the photoelectric effect, radioactivity or atomic spectra. On the subject test there are also occasional questions on the history of science, which basically concern matching names with discoveries. For this reason, we have mentioned the discoverers of the various phenomena through the text. The chronology is summarized in Section 10. We begin, however with basic atomic and nuclear structure because at the level of the SAT exams quantum mechanics is unnecessary. Although there is an excess of terminology in atomic and nuclear physics, most is self-evident and in the way of math little more than arithmetic is required.

2. Basic Atomic Physics

As you certainly know, an **atom** consists of a **nucleus** surrounded by **electrons**. The nucleus itself is composed of **protons** and **neutrons**. Because the number of positively charged protons in the nucleus equals the number of negatively charged electrons orbiting around it, an atom is ordinarily neutral. The neutron itself, as its name implies, is an electrically neutral particle. Both protons and neutrons are about 1600 times as massive as electrons, which means that virtually the entire mass of an atom resides in the nucleus.

The number of protons in the nucleus is referred to as the **atomic number.** The atomic number is generally designated Z. Because the proton carries a positive charge e equal in magnitude to the charge on the electron, Ze gives the number of positive charges in the nucleus. Atomic number is what distinguishes one chemical element from another on the **periodic table**. For example, hydrogen, the lightest element has an atomic number of 1, meaning that its nucleus consists of a single proton. Helium, with an atomic number of 2, has two protons in its nucleus and so forth.

In the early twentieth century it was generally believed that the electrons and protons were spread out evenly over the space occupied by an atom. The existence of a nucleus was established only in 1911 through experiments carried out by Hans Geiger and Ernest Marsden under the direction of Ernest Rutherford. Geiger and Marsden bombarded thin gold foil with helium nuclei, also termed **alpha particles**, which were emitted by a decaying radioactive source. The alpha particles usually passed through the foil, but occasionally bounced straight back, indicating that an atom consisted mostly of empty space, with a small nucleus at the center. The model Rutherford developed is the picture of the atom we frequently imagine today: a nucleus surrounded by electrons orbiting at large distances.

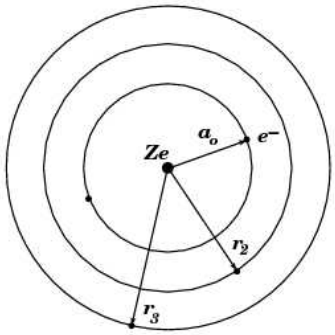


Figure 12.1

Figure 12.1 illustrates the **Rutherford atom** with the nucleus Ze at the center and several possible electron orbits. Rutherford was able to estimate that the size of the nucleus was approximately 10^{-14} m (Example 12.1). The size of the smallest electron orbit for hydrogen, a_o, is called the **Bohr radius** and is about 10^{-10} m. (We derive the Bohr radius in Section 8.) Therefore, **the size of an atom is about 10,000 times larger than the size of the nucleus.** If an atom were the size of a football stadium, its nucleus would be about the width of an icon on your smart phone.

The electrons in an atom are not allowed to orbit the nucleus at arbitrary radii. Due to the basic postulates of **quantum mechanics** (Section 6**)** electrons can occupy only discrete **energy levels**, which are associated with specific orbital radii. For example, it was known since the nineteenth century that hydrogen emits light in certain **spectral lines** spaced at uneven wavelength intervals, as in Figure 12.2.

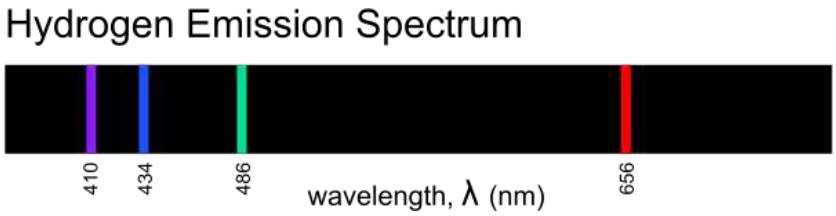


Figure 12.2

According to quantum mechanics, this light is emitted when an electron drops from one atomic energy level to another, emitting a **photon**, or quantum of light. Like particles, photons have a momentum, but they also have a frequency and wavelength, like a wave. In this case the frequency is proportional to the energy difference between the two levels; the greater the difference, the greater the photon frequency and energy. Because the energy levels in a hydrogen atom are unevenly spaced, the emitted frequencies and wavelengths are also spaced at uneven intervals, as the illustration shows.

Since the energy levels for real atoms may be difficult to calculate, for the purpose of exams one often makes use of imaginary systems. For example, Figure 12.3 shows a hypothetical atom in which the energy levels are equally spaced by 10 electron volts.

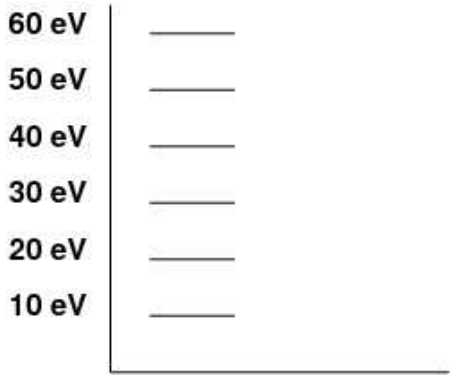


Figure 12.3

The eV is a convenient energy unit for quantum problems because atomic systems, where quantum effects frequently take place, often have energies in the 10 − eV range. This imaginary system can only absorb or emit photons whose energy occurs in integral multiples of 10 eV. When the system absorbs a photon, an electron will jump to a higher energy level. We say that the electron has been **excited**. Normally, electrons in an excited state quickly decay to a lower energy state, in the process emitting a photon (usually denoted by γ) whose energy equals the energy difference between the two states (Figure 12.4). Such a process is called **spontaneous emission** and produces the spectral lines in Figure 12.2. Because the transition between levels occurs in discrete amounts, they are generally called **quantum jumps.** The lowest possible energy level of a system is termed the **ground state**.

What a quantum system *cannot* do is make transitions to energies lying between the levels (dashed arrows in Figure 12.4).

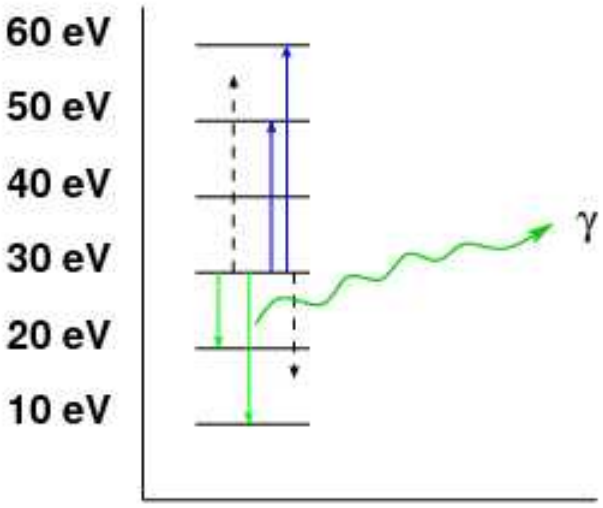


Figure 12.4

Exercise: Calculate the frequency corresponding to each wavelength shown in Figure 12.2.

3. Nuclear Basics

As mentioned, the atomic number Z gives the number of protons in a nucleus, which defines a chemical element. Most elements, however, are present in several **isotopes**, which differ from one another in the number of neutrons they contain. Hydrogen comes in three isotopes: ordinary hydrogen (one proton, no neutrons); deuterium, or heavy hydrogen (one proton, one neutron); and tritium, super heavy hydrogen (one proton, two neutrons).

The number of neutrons in an isotope is ordinarily denoted by N. The total number of neutrons and protons together in the nucleus is termed the **mass number** and is generally designated A. Thus, $A = Z + N$. Neutrons and protons together are collectively termed **nucleons** and so A equals the number of nucleons in the nucleus.

The atomic number Z usually appears as a subscript to the left of the abbreviation for the element and the mass number A as a superscript. Thus, ${}_1^1\mathrm{H}, {}_1^2\mathrm{H}, {}_1^3\mathrm{H}$ represent ordinary hydrogen, deuterium and tritium, respectively. Often the subscripts are omitted, so ${}^1\mathrm{H}$, ${}^2\mathrm{H}$ and ${}^3\mathrm{H}$ represent the same thing. When discussing nuclei, as opposed to atoms, the term **nuclide** is frequently used instead of element. A carbon-14 nuclide is the nucleus of the carbon isotope containing 6 protons and 8 neutrons.

Because protons are positively charged, they repel each other and would tear the nucleus apart if another force did not bind the nucleons together. This force is termed the **strong nuclear force** and it acts with equal strength on neutrons and protons. Like the Coulomb force, the strong nuclear force is a fundamental force of nature and it is about 100 times stronger than the electric repulsion between protons. It is, however, an extremely short-range force, operating only within the atomic nucleus, which is why electromagnetic forces dominate in daily life where distances are much larger. Heavier isotopes tend to contain more neutrons than protons. If only protons were present, one could not even assemble a nucleus; the presence of neutrons allows the strong force to counteract the Coulomb repulsion, making the nucleus more **stable**. Because the Coulomb repulsion is long-range but the strong force is short-range, as the nucleus gets larger, one requires an excess of neutrons to counteract the Coulomb force. The AP exam expects you to be familiar with this concept.

Example 12.1: Estimating the size of the nucleus. Geiger and Marsden fired alpha particles at thin foils of gold. Alphas consist of 2 neutrons and 2 protons. A gold nucleus, with atomic number 79, contains 79 protons. The gold and alphas therefore repel each other. In order for the alphas to bounce off the gold nuclei, they must first be brought to a complete stop by the Coulomb force between the protons. Geiger and Marsden's alpha particles had an initial kinetic energy $1/2\, m_\alpha v_o^2$ of about 10^{-12} J, where v_o was the initial velocity.

The electrostatic force that the gold nucleus exerts on the alphas is $F = m_\alpha a = k(2e)(79e)/r^2$, or $a \approx 160ke^2/m_\alpha r^2$. From the basic kinematic formula, $v_o^2 - v_f^2 = 2ad$; at the point where the alphas bounce, $v_f = 0$. Thus, $v_o^2 = 2ad$, where d is the distance over which the alphas are stopped. From the initial kinetic energy, $v_o^2 = 2 \times 10^{-12}/m_\alpha$. Then

$$2 \times \frac{10^{-12}}{m_\alpha} \approx 2 \times 160 \frac{ke^2}{r^2 m_\alpha} \times d\,.$$

Notice that the 2's and the m_α's cancel. The stopping distance d must be something like the size of the nucleus, since F is strongest near the nucleus and decreases very rapidly away from it. For an estimate we can then set $d \approx r$ and solve for r:

$$r \approx \frac{160ke^2}{10^{-12}}.$$

Plugging in the numbers gives $r \approx 3.6 \times 10^{-14}$ m, about twice the size of observations.

This problem can be done more simply at the AP level by realizing that at the moment the alphas bounce off the gold, all the energy is potential: $U = -k(2e)(79e)/R$, where R is the radius of the nucleus. Therefore, by conservation of energy

$$\frac{1}{2}mv_o^2 = \frac{158ke^2}{R}.$$

Solving for R and using 10^{-12} J for the kinetic energy gives the same result as above.

4. Radioactivity

Most nuclides are **radioactive**: they spontaneously emit a particle and in the process decay from one nuclear species into another, in the process releasing large amounts of energy. Although textbooks credit Henri Becquerel with the discovery of radioactivity in uranium in 1896, Claude Felix Abel Niepce de Saint-Victor published numerous memoirs on similar experiments forty years earlier. Marie and Pierre Curie also carried out pioneering work in radioactivity, discovering radium and polonium.

There are many different types of radioactive decay. Ernest Rutherford determined that the known radioactive isotopes gave off three types of radiation, which differed by their ability to penetrate material. Not knowing their true identities, he labeled them simply **alpha**, **beta** and **gamma**. Later, Rutherford and others determined that alpha radiation was the emission of a helium nucleus, beta the emission of an electron and gamma the emission of light. Nevertheless, the original terminology is still routine today, in particular on exams. For example, we say that uranium-238 decays into thorium-234 through alpha emission. The actual reaction is $^{238}\text{U} \rightarrow {}^{234}\text{Th} + {}^{4}\text{He}$. When not bound into an atomic nucleus, the neutron is a radioactive particle, which decays into a proton through beta emission: $n \rightarrow p + e^- + \overline{\nu}$. The $\overline{\nu}$ represents a **neutrino** (technically an antineutrino), which as its name implies is a neutral particle whose mass until recently was believed to be zero. Today the neutrino is known to have a tiny mass, about one millionth that of an electron.

The rate at which a given isotope decays can be characterized by its **radioactive half-life**.

The half-life of a radioisotope is the time during which half the nuclei of that isotope in a given sample decay.

Half-lives vary by many orders of magnitude from one isotope to another. The half-life of ^{235}U is about 700 million years, while the half-life of ^{238}U is about 4.5 billion years. The half-life of tritium is 12.3 years. The half-life of a free neutron is about ten minutes. The longer the half-life, the lower the decay rate.

By definition, after one half-life, half the atoms in a radioactive sample have decayed away. After two half-lives, half of the remainder have decayed away, leaving one-quarter of the original, after three half-lives an eighth, and so forth. **The half-life of an element is designated $t_{1/2}$.** If N_o is the original number of atoms in the sample and N is the number remaining after a time t, then

$$\frac{N}{N_o} = 2^{-\frac{t}{t_{1/2}}} \quad (12.1)$$

After one half-life $t/t_{1/2} = 1$, which gives $N/N_o = 2^{-1} = 1/2$, as just stated. After two half-lives $t/t_{1/2} = 2$, giving $N/N_o = 2^{-2} = 1/4$. After three half-lives $N/N_o = 2^{-3} = 1/8$, etc. If you find it easier, **you can write $N/N_o = 1/2^n$, where n is the number of half-lives (n does not need to be an integer).**

Example 12.2: Age of the earth. The ratio of ^{235}U to ^{238}U in mineral deposits is typically about 0.7 percent, or 0.007. Assume that at the creation of the earth the two isotopes were present in equal amounts. Call the initial amount of each 1. From above, 4.5 billion years is one-half life of ^{238}U and 6.4 half-lives of ^{235}U. By definition, after 4.5 billion years ^{238}U will have decayed to half its original amount, 0.5. In the same period ^{235}U will have decayed to $2^{-6.4} = 0.012$. At that time the ratio of ^{235}U to ^{238}U is $0.012/0.5 = 0.024$, a factor of about 3 above what is measured. That suggests letting ^{235}U decay for two more half-lives. After eight half-lives, the amount of ^{235}U remaining will be $2^{-8} = 0.0039$. But eight half-lives of ^{235}U is 1.2 half-lives of ^{238}U, so in that period ^{238}U has decayed to an amount $2^{-1.2} = 0.435$ Then we have $^{235}\text{U}/\,^{238}\text{U} = 0.0039/0.435 = 0.009$, slightly higher than what is measured. If the guess that the two isotopes were formed in equal amounts is correct, it indicates that the age of the earth is approximately 8 half-lives of ^{235}U, or 5.6 billion years.

If one knows the **decay chain** of the radioisotope, one can do the above type of calculation more convincingly. Suppose that an isotope's original nuclei decay into a known product p. For example, when we say ^{238}U decays with a half-life of 4.5 billion years, this actually represents a long series of decays whose final product is ^{206}Pb (lead). Solving Eq. (12.1) for t gives

$$t = -\frac{t_{1/2}}{\ln 2} \ln\left(\frac{N}{N_o}\right).$$

However, at any time the number of the original species remaining (N) plus the number of product atoms (N_p) must equal the initial number of atoms (N_o). Inserting $N_o = N + N_p$ into the above equation gives for the age of the sample

$$t = \frac{t_{1/2}}{\ln 2} \ln\left(1 + \frac{N_p}{N}\right).$$

If in mineral deposits $^{206}\text{Pb}/\,^{238}\text{U} = 0.65$, we set $N_p/N = 0.65$. Letting $t_{1/2} = 4.5 \times 10^9$ yr gives $t = 3.25 \times 10^9$ yr for the age of this deposit.

5. Energy in Nuclear Reactions

In the radioactive process $^{238}\text{U} \rightarrow {}^{234}\text{Th} + {}^{4}\text{He}$, a single parent nuclide, uranium-238, decays into two daughter nuclides, thorium-234 and helium-4. During the decay, the system moves from a higher energy state to a lower energy state, like a ball rolling down a hill. Energy is therefore released. In almost every situation we wish to determine the amount, which for nuclear processes turns out to be a million times larger than for ordinary chemical reactions.

To determine the energy released in a nuclear reaction, we introduce the **atomic mass unit**, the traditional quantity for working nuclear physics problems. The atomic mass unit (u) is defined such that the *atomic mass* of one carbon-12 atom is exactly 12 u. Thus,

$$1\text{ u} = 1.6605 \times 10^{-27}\text{ kg}$$

However, for nuclear physics the kg is an extremely inconvenient unit, being about 10^{27} times larger than a nuclear mass. It is far more sensible to use MeV, or million electron volts. From Section 9.13 $1\text{ eV} = 1.6 \times 10^{-19}$ J; therefore, $1\text{MeV} = 1.6 \times 10^{-13}$ J. The MeV, though, is a unit of energy, whereas the kg is a unit of mass. To convert kilograms to energy we use Einstein's famous equation $E = mc^2$ (Section 6.4). Multiplying the above by c^2 gives $1\text{ u} = 1.492 \times 10^{-10}$ J and converting to MeV shows that

$$1\text{ u} = 931.5\text{ MeV}$$

This is a good number for your number bank—and an even better one is to remember that **the mass of a proton or neutron is very nearly 1 u.**

We now examine the reaction $^{238}\text{U} \rightarrow {}^{234}\text{Th} + {}^{4}\text{He}$. In terms of atomic mass units, $^{238}\text{U} = 238.0508$; $^{234}\text{Th} = 234.0436$; $^{4}\text{He} = 4.0026$. Adding together the mass of the daughter particles gives 238.0462, which is *less* than the mass of the parent, ^{238}U. The difference $\Delta m = 0.0046$ u is termed the **mass deficit.** We are reluctant to believe that mass has disappeared in this reaction and so are forced to extend the separate laws of conservation of mass and energy to

Conservation of mass-energy:

The total mass-energy of an isolated system remains constant.

In any reaction the sum of all the energies must balance, including energy associated with mass by $E = mc^2$. By the same token, the sum of all the masses must balance, where mass includes the mass associated with energy by $m = E/c^2$. The current equation balances if we assume that an energy equivalent to the mass deficit has been released: $m_U c^2 = m_{Th}\, c^2 + m_{He} c^2 + \Delta m c^2$. **The value of the energy deficit is $\Delta E = \Delta m c^2$.** In this case, $\Delta E = 4.2849$ MeV.

Example 12.3: Fission. When a slow neutron hits a ^{235}U nucleus, it can render the uranium-235 unstable and cause it to fragment into two lighter pieces plus several neutrons, as shown in Figure 12.5.

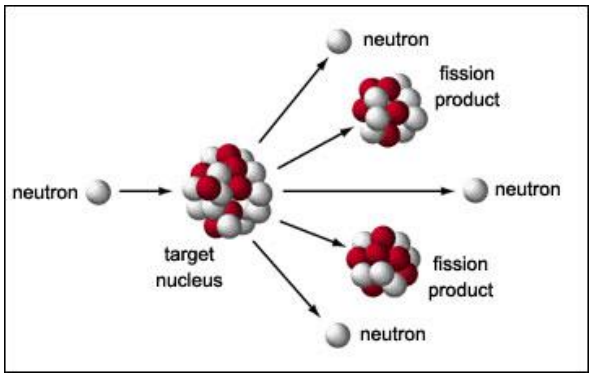


Figure 12.5

Such a processes is termed **fission**. There are many possible fission reactions. One is the fissioning of uranium-235 into rubidium and cesium: $\text{n} + {}^{235}\text{U} \rightarrow {}^{90}\text{Rb} + {}^{143}\text{Cs} + 3\text{n}$. In terms of atomic mass units, $n = 1.00866$ and $^{235}\text{U} = 235.0429$. Thus, the total mass of the reactants is $m_o = 236.0516$ u. On the right-hand-side of the equation, $^{90}\text{Rb} = 89.9148$ and $^{143}\text{Cs} = 142.9273$. Thus, the total mass of the products is $m_f = 235.8680$ u, which is less than m_o. The mass deficit is then $\Delta m = 0.184$ u and the energy released is $\Delta E = \Delta m c^2 = 171\text{ MeV} = 27.4 \times 10^{-12}$ J.

Now, the mass of a ^{235}U atom is 3.9×10^{-25} kg. Thus, a kg of uranium contains about 2.5×10^{24} atoms and if they all fission, the total energy released is 7×10^{13} J. By contrast, a kg of gasoline releases about 5×10^{7} J. We see that fission reactions are a million times more efficient than chemical reactions.

The energy released in a nuclear reaction is also referred to as its **Q value**. Where does the released energy come from? It is the energy associated with the strong nuclear force that holds the nucleus together and is referred to as the nuclear **binding energy**. You can view the binding energy either as the energy that is released when you rip a nucleus into its constituent nucleons, or as the energy necessary to assemble the constituents into a nucleus. In this sense it is totally analogous to the electrical potential released when a charge was removed from the triangle in Example 9.5. The Q of a reaction is the difference in binding energy between the reactants and products.

Consider the **fusion** reaction ${}_1^2\mathrm{H} + {}_1^3\mathrm{H} \rightarrow {}_2^4\mathrm{He} + \mathrm{n}$. As opposed to a fission reaction, here deuterium *combines* with tritium to produce helium plus a neutron. In terms of atomic mass units, ${}_1^2\mathrm{H} = 2.041$; ${}_1^3\mathrm{H} = 3.016$. The total mass of the reactants is $m_o = 5.030$ u and the mass of the products is $m_f = 5.011$ u. Therefore, the mass deficit is $\Delta m = 0.0187$ u and the energy released is $Q = (.0187) \times (931.5) = 17.5$ MeV. Again, this energy is the difference in binding energy between the reactants and products. Of the released energy, 14.1 MeV is carried off by the neutron and 3.5 MeV by the helium.

Although fission produces approximately 10 times more energy per reaction than fusion, 1 kg of deuterium or tritium contains about 100 times more molecules than 1 kg of uranium. Therefore, on a per kilogram basis, fusion releases roughly ten times the energy of fission.

Example 12.4: Who gets the energy? In the fusion reaction just discussed the neutron gets about 4 times as much energy as the helium-4 nucleus. The reason is conservation of momentum. If deuterium and tritium collide head on, they have roughly equal and opposite momenta and the system's total momentum is approximately zero. The neutron and helium-4 nucleus must then also have equal and opposite momenta: $m_n v_n + m_{He} v_{He} \approx 0$. Since $m_{He} = 4m_n$, the neutron is released with four times the velocity of the helium-4 nucleus. The ratio of their kinetic energies is

$$\frac{KE_n}{KE_{He}} = \frac{m_n v_n^2}{m_{He} v_{He}^2} \approx \frac{1}{4} \times 16 = 4,$$

which is very nearly the observed ratio $14.1/3.5$.

You may object that tritium is more massive than deuterium; therefore, they do not have equal and opposite momenta. True, but the difference is not as great as it seems. If the two nuclei are at equal temperatures in a gas, then Eq. (8.8) tells us that they have the same kinetic energy, or $m_D v_D^2 = m_T v_T^2$. But $m_T = \frac{3}{2} m_D$, implying that $v_T = \sqrt{2/3}\, v_D$ and the momentum of the tritium is only $\sqrt{3/2} = 1.22$ times the momentum of the deuterium.

Note that in both the fission and fusion reactions we considered, the mass of the products was *less* than the mass of the constituents, leading to the release of binding energy. If one plots the binding energy per nucleon versus atomic mass one gets Figure 12.6

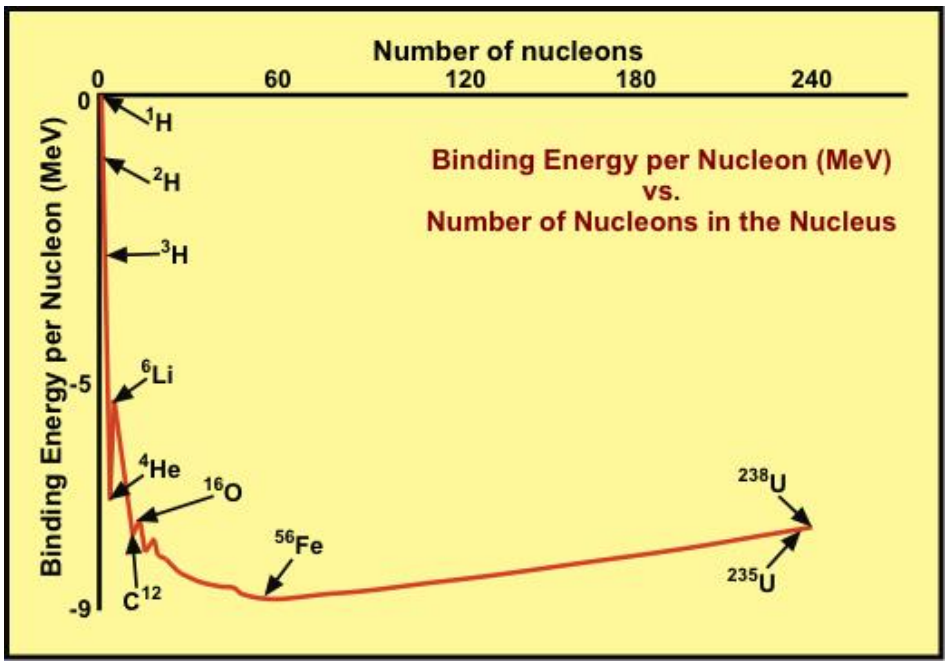


Figure 12.6

This is the famous "curve of the binding energy." For light nuclides, the magnitude of the binding energy per nucleon becomes greater[1] as the atomic mass increases until ^{56}Fe, iron. Then it slowly decreases toward uranium. The curve tells us that iron is the most tightly bound, or stable, element. Thus, reactions "roll down the hill" to iron and release energy. To the left of iron, combining light nuclei releases energy. Fusion. To the right of iron, breaking heavy nuclei into two lighter, more stable fragments releases energy. Fission.

Exercise: Convince yourself that the values for binding energy on the curve are consistent with the energy released that we found for ${}_1^2\mathrm{H} + {}_1^3\mathrm{H} \rightarrow {}_2^4\mathrm{He} + \mathrm{n}$.

6. The Quantum Hypothesis

Until 1900 physicists believed that energy, whether potential or kinetic, could come in arbitrary values. In particular, you could decrease the energy of a mechanical system continuously until it reached zero. In 1900 Max Planck was attempting to understand the **spectrum** of **blackbody radiation**.

[1] Like gravitational potential energy, the binding energy is negative. It becomes more negative as one goes from hydrogen to helium to iron. On many graphs the negative sign is omitted and the curve is presented upside down from this one, which obscures the fact that energy is being *released* toward iron.

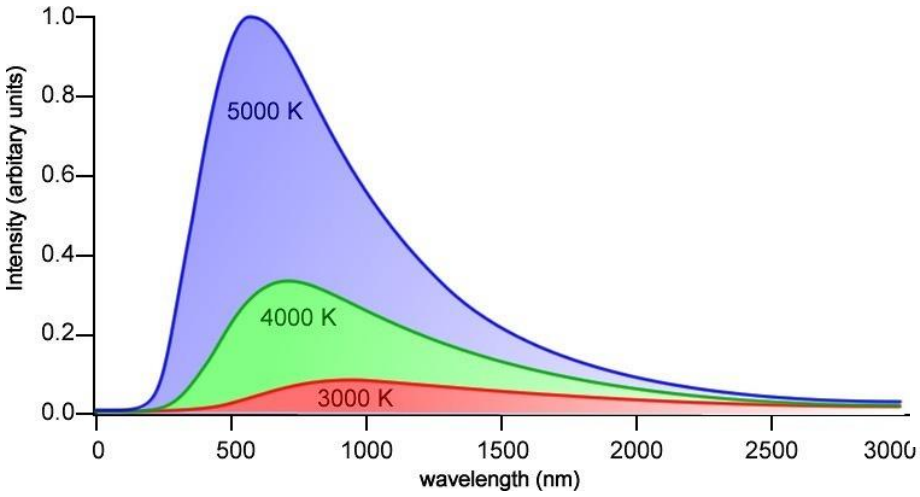


Figure 12.7

Blackbody radiation is physics-speak for the thermal radiation given off by any hot body, for example a glowing metal iron in a fireplace. Thermal radiation consists of ordinary incoherent electromagnetic waves (Section 11.10); however, at room temperatures the wavelengths are in the infrared part of the spectrum, which is invisible to the eye, and we refer to it as heat. The spectrum of an object is a graph of the intensity of radiation emitted in a given frequency (or wavelength) interval. Intensity is defined as power per unit area (Section 11.10), so a spectrum is generally the emitted power per unit area per unit frequency. Figure 12.7 shows a blackbody spectrum, which in the ideal case depends only on the temperature of the object.

Classical physics was unable to account for the shape of the blackbody curve. It was known that oscillating electric charges emitted electromagnetic radiation, or light. Planck explained the blackbody curve by assuming that the hot object was composed of mysterious oscillators that radiated light at their oscillation frequency. (Planck never explained exactly what the oscillators were.) However, he was forced to propose that the energy of the emitted light had a minimum value given by

$$E = h\nu \quad (12.2)$$

where ν ("nu") was the frequency of the oscillator[2] and h was a new fundamental constant

$$h = 6.63 \times 10^{-34}\ \text{kg m}^2/\text{s}$$

Today h is universally referred to as **Planck's constant**. Notice that the dimensions of h are energy × time. These dimensions are shared by no other constant we have previously encountered, which indicates that new physics is involved.

If a problem contains h, quantum mechanics is involved; if it does not contain h, quantum mechanics is not involved.

Planck also made a second assumption, that the oscillators could only emit or absorb light with energy

$$E = nh\nu \quad (12.3)$$

[2] You can write $E = hf$ if you prefer, but physicists virtually always use $E = h\nu$.

where n is an integer greater than or equal to 1. Notice that in Eq. (12.2) the minimum energy varies continuously with frequency and so can take on arbitrary values, but once the minimum is established for a given frequency, Eq. (12.3) requires that the energy of light emitted by a body must be an *integral* multiple of that minimum. The allowed energies of a system, designated by the **principal quantum number** n, are the **energy levels** discussed in Section 2. In Planck's blackbody the energy levels are spaced by equal intervals $\Delta E = h\nu$ as in the hypothetical system of Figure 12.3. In that example electrons can jump only between levels spaced by multiples of $10\text{ eV} = h\nu$

Eqs. (12.2) and (12.3) form the basis of all quantum mechanics. **Nail them.**

The reason we do not often observe quantum effects in daily life is that Planck's constant is an extremely small number, as the following illustrates.

Example 12.5: A Planckian spring. A spring of mass $m = 100\text{ gm}$ and spring constant $k = 150\text{ N/m}$ is subject to the rules of quantum mechanics. What are the allowed spacings of the energy levels of the spring?

From Eq. (5.11), the natural angular frequency of the spring is $\omega = \sqrt{k/m} = 38.7\text{ rad/s}$. Then $\nu = \omega/2\pi = 6.16\text{ s}^{-1}$. From Eq. (12.3) the allowed energies of the spring are $E = nh\nu$ and the spacing of the energy levels ΔE for successive integers is simply $h\nu$, as above. Thus, $\Delta E = 6.16\,h \approx 4 \times 10^{-33}$ J.

The result is so extraordinarily small that for all intents and purposes the spacing between energy levels is zero and the allowed energies of the spring may be regarded as continuous. This is why you do not often observe quantum effects in daily life. On the other hand, if the spring had the mass of an electron, $m = 10^{-30}$ kg, the same calculation gives $\Delta E \approx 1.3 \times 10^{-18}$ J, which is in the $10 - \text{eV}$ range and can be easily observed.

7. The Photoelectric Effect and Photons

The photoelectric effect was discovered by Heinrich Hertz and several other investigators in 1887.

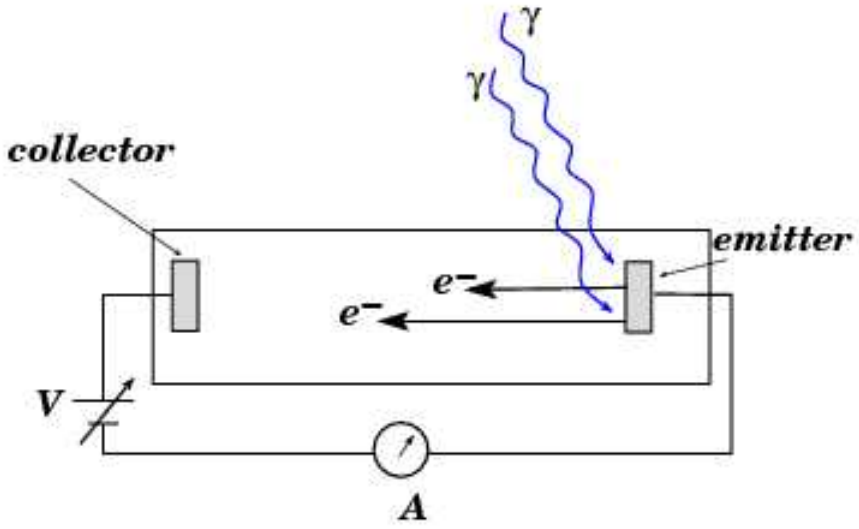


Figure 12.8

The effect is easy to describe but impossible to explain by classical physics. As illustrated in Figure 12.8, a light beam (γ) was incident on piece of metal. Regardless of the intensity of the beam, if the light was below a certain frequency, termed the **threshold frequency**, nothing whatsoever happened. Above the threshold frequency, **photoelectrons** (e^-) were emitted by the metal. The photoelectrons traveled to a collector, producing a small electric current that could be measured by a simple ammeter (A). However, once above the threshold frequency, increasing the intensity of the beam did *nothing* to alter the energy of the electrons. On the other hand, increasing the frequency of the light *did* increase the electron energy. (The electron energy was measured by adjusting a voltage source (V) so that at a voltage called the **stopping potential**, the current was reduced to zero. At the stopping potential, eV must equal the electrons' kinetic energy. Nevertheless, increasing the intensity of the beam did *not* increase the required stopping potential, indicating that the intensity was unrelated to the electron energy.)

The photoelectric effect was so puzzling because according to classical physics, the intensity of a light beam does measure its energy. Yet, no matter how intense the beam, if the light was below the threshold frequency, no electrons were ejected from the metal. Above the threshold, one would expect a brighter beam to produce more energetic electrons. However, as just stated, increasing the intensity *failed to* increase the energy of the electrons, yet it *did* increase the *number* of emitted electrons (the electrical current). On the other hand, according to classical physics, the energy in a light beam does *not* depend on its frequency. But increasing the light frequency above the threshold *did* increase the electron energy.

If you are not baffled by all this, you should be. In 1905 Einstein explained away all the contradictions in the photoelectric effect with a simple but radical hypothesis:

1. Light was composed of particle-like quanta whose energy was given by Planck's formula $E_\gamma = h\nu$.

2. The energy of the emitted photoelectrons was given by

$$E_e = E_\gamma - W \quad (12.4)$$

where W was the **work function** of the particular metal, the minimum energy required to knock out an electron.

3. Each light quantum knocked a single electron out of the metal.

Postulate 1 differed from Planck's in that Planck never said that light came in the form of particles; he said only that the energy of light was $E_\gamma = h\nu$. According to Einstein's postulate 2, a graph of the photoelectron energy versus the energy of the incoming light quanta was just a straight line, as in Figure 12.9.

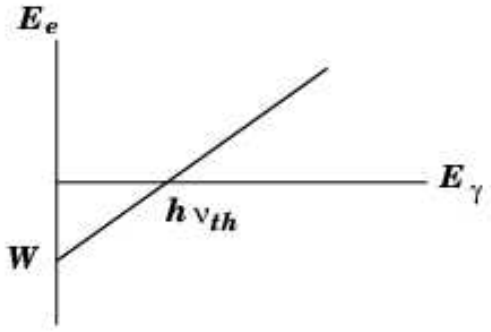


Figure 12.9

We see that below a threshold frequency ν_{th} the electrons do not have enough energy to escape the metal and no photoelectrons are emitted, as was observed in the experiment. Above ν_{th}, the energy of the photoelectrons increases linearly with ν, as was also observed. **We interpret the intensity of the light beam as proportional to the number of light quanta**. If each light particle knocks out a single electron, increasing the intensity of the beam should cause more electrons to be emitted, as was again observed. Thus, Einstein's picture accounted for all the observations associated with the photoelectric effect.

About two decades after he made the proposal, Einstein's light quanta, or particles, became known as **photons**. Photons, as far as is known, have no mass, but the fact that they behave like particles implies that they carry not only energy but momentum. In Newtonian physics the momentum p of a particle is given by mv and its kinetic energy by $1/2\,mv^2$. Since photons travel with velocity c one might reasonably conclude that the relationship between the energy and momentum of a photon should be $E = pc/2$. *This is incorrect*. We state without proof that the momentum of a photon is given by

$$p = \frac{E}{c} \quad (12.5)$$

where $E = h\nu$. This requirement to know this fact occasionally surfaces on an AP exam.

Example 12.6: Photoelectric effect with zinc. Zinc is the most common metal employed in demonstrating the photoelectric effect. Its work function is 4.3 eV. What is the threshold frequency for the emission of photoelectrons? What is the threshold wavelength? If the stopping potential is 8 V, what is the maximum velocity of the electrons?

Einstein's equation for the photoelectric effect is $E_e = E_\gamma - W$, or $E_e = h\nu - W$. By definition, $E_e = 0$ at the threshold frequency. Thus, $h\nu_{th} = W$.

Tip: Planck's constant has the value 4.14×10^{-15} eV s.

If $W = 4.3$ eV, then $\nu_{th} = 1.04 \times 10^{15}$ Hz. With $c = \lambda\nu$, we get $\lambda_{th} = 288$ nm, which is in the ultraviolet.

The stopping potential is the voltage that reduces the photoelectron current to zero. The electric field between the collector and emitter must therefore be just sufficient to prevent the electrons from reaching the collector. In that case $eV = 1/2\,mv^2$ (problems 9.34-9.35). Solving for v with the mass and charge of the electron and $V = 8$ V, gives $v \approx 1.7 \times 10^6$ m/s.

8. Atomic Spectra

One of the outstanding puzzles in the late nineteenth and early twentieth centuries was to explain the spectrum of light emitted by atomic hydrogen, the same spectrum displayed in Figure 12.2. A hot body produces radiation at all frequencies; its spectrum is said to be *continuous*. By contrast, when an electric current is passed through a glass tube containing hydrogen gas the result is the discrete series of lines in Figure 12.2. What's more, the frequencies or wavelengths of the lines are spaced in no obvious pattern, which made the hydrogen spectrum virtually impossible to understand.

In 1885 Jacob Balmer, a numerologist, showed by trial and error that the visible wavelengths of the hydrogen spectrum were given exactly by a formula that can be written

$$\frac{1}{\lambda} = R\left(\frac{1}{2^2} - \frac{1}{n^2}\right) \quad n = 3,4,5 \ldots \quad (12.6)$$

The number $R = 1.097 \times 10^7\ \text{m}^{-1}$ is called the **Rydberg constant**. Once the blackbody spectrum and photoelectric effect had been accounted for, to give a physical explanation of the **Balmer formula** (12.6) became the next great challenge in physics.

The dimensions of Planck's constant are energy × time $= mv^2 \times t$. If you think about it for a moment, this is the same as $mv \times vt = mvr$, or angular momentum.

The dimensions of Planck's constant are angular momentum.

In 1912 John Nicholson suggested that the angular momentum of electrons in atoms was quantized and could change only in integral multiples of Planck's constant. With his model of the atom, however, he could not reproduce the Balmer formula. In 1913 Niels Bohr took up Nicholson's suggestion and applied it to a Rutherford-model hydrogen atom (Figure 12.10a), in which a single electron orbits a single proton.

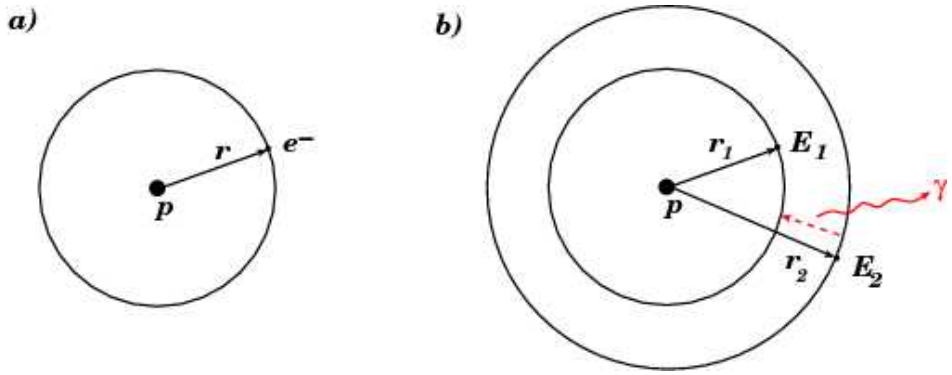


Figure 12.10

The Rutherford model, however, brings with it several difficulties, one of which is that accelerating charges radiate energy. Because an electron is in a circular orbit around the proton, it is undergoing centripetal acceleration and must lose energy. Calculations show that the electron should spiral into the nucleus in less than a billionth of a second. The Rutherford atom was therefore unstable. A reality check, however, confirms that atoms exist for more than a billionth of a second. To overcome this and similar dilemmas, Bohr postulated:

1. Following Nicholson, the angular momentum L of the electrons is given by $L = mvr = n\hbar$ where $\hbar$ ("h-bar") is the universal symbol for $h/2\pi$ and n is an integer. No other values for L are permitted.

2. When an electron jumps from one orbit to another it absorbs or emits a photon (γ) with energy equal to the energy difference between the two orbits: $\Delta E = E_2 - E_1 = h\nu$. (Figure 12.10b shows the emission of a photon.)

3. Unless an electron changes orbits, it does not radiate or absorb energy. The allowed orbits are stable.

All three postulates contradicted classical physics. With them, however, it was not difficult for Bohr to reproduce the Balmer formula. Assume any hydrogen-like atom consists of an electron with charge $-e$ in a circular orbit around the nucleus, which has charge $+Ze$. The electrostatic potential energy of the electron is $U = -kZe^2/r$ and its total energy $E = KE + U$ becomes

$$E = \frac{1}{2}mv^2 - \frac{kZe^2}{r}.$$

However, because the electron is in a circular orbit, it experiences a centripetal force equal to the Coulomb force acting on it (problem 5.25). We must then have

$$\frac{mv^2}{r} = \frac{kZe^2}{r^2} \qquad (12.7)$$

implying that $1/2\, mv^2 = kZe^2/2r$. Consequently, the total energy is just one-half the potential energy:

$$E = -\frac{1}{2}\frac{kZe^2}{r} \quad (12.8)$$

We also have from Eq. (12.7)

$$v = \sqrt{\frac{Zke^2}{mr}}.$$

The electron's angular momentum is then $L = mvr = \sqrt{mrZke^2}$, and $L^2 = mrZke^2$. By postulate 1, this gives $L^2 = n^2\hbar^2$ and so

$$r_n = \frac{n^2\hbar^2}{mZke^2} \equiv n^2 a_o\,,$$

where r_n refers to the orbital radius corresponding to the given value of n and $a_o \equiv \hbar^2/mZke^2$. Working out this collection of constants when $Z = 1$ (**Exercise**) gives

$$a_o = 5.29 \times 10^{-11}\ \text{m}$$

The length a_o, mentioned in Section 2, is termed the **Bohr radius** and it gives the radius of the smallest possible orbit for an electron around the proton; it is essentially the size of a hydrogen atom. Substituting the expression for r_n into Eq. (12.8) yields the energy of the nth orbit:

$$E_n = -\frac{1}{2}\frac{mk^2Z^2e^4}{n^2\hbar^2} \quad (12.9).$$

By postulate 2, the difference in energy between orbits is $E_{n2} - E_{n1} = h\nu$. The frequency is $\nu = c/\lambda$, and so

$$\frac{1}{\lambda} = \frac{1}{2}\frac{(2\pi)^2 mk^2Z^2e^4}{ch^3}\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right) \quad (12.10)$$

with $n_2 > n_1$. Working out the collection of constants for $Z = 1$ (**Exercise**) shows that they exactly equal the Rydberg constant R in Eq. (12.6). Thus, Bohr fully accounted for the Balmer formula when $n_1 = 2$. When $n_1 = 1$ you get a different series of lines, called the **Lyman series**. When $n_1 = 3$ you get the **Paschen series**. The Bohr atom is a good example of a system in which the energy levels are *not* evenly spaced but are proportional to $1/n^2$.

Notice that as n_2 gets very large, $1/n_2^2$ becomes very small. In the limit $n_2 \to \infty$ the energy difference between the two levels is just E_n in Eq. (12.9), where now $n = n_1$. The most negative E_n corresponds to $n_1 = 1$. Then working out the constants gives $E = -13.6$ eV. This is the **ground state** of hydrogen, the lowest energy level. By the same token, it takes $+13.6$ eV to remove an inner electron from a hydrogen atom, or to **ionize** the hydrogen. The amount of energy required to ionize an atom is termed its **ionization potential**.

Exercise: Find the wavelengths in the Balmer series for $n_1 = 2$ and $n_2 = 3, 4, 5$. Find the wavelengths for the Lyman series: $n_1 = 1$ and $n_2 = 2, 3, 4$.

9. Matter Waves

The photoelectric effect implied that light has particle properties. In 1924 Louis de Broglie proposed that matter has wave properties. Specifically, he conjectured that particles have associated with them a wavelength given by

$$\lambda = \frac{h}{p} \quad (12.11)$$

where p is the particle's momentum. For particles moving at much less than the speed of light, p is just the Newtonian expression mv. The wavelength given by Eq. (12.11) is referred to as the **de Broglie wavelength** of a particle. For an electron moving at $v = 10^6$ m/s, $\lambda = 0.73$ nm. This is a very small number but measureable. Indeed, within a few years after de Broglie made his proposal the existence of **matter waves** was verified by experiments showing that electrons could be diffracted in the same way as light waves (below) and that they must therefore have wave properties.

Tip: The de Broglie wavelength can be remembered by realizing that dimensionally the only way to form a wavelength from Planck's constant is to divide it by a momentum.

We have $[h] = \text{energy} \times \text{time} = ML^2/T$. We want a wavelength, which is a length, L. Thus, we must divide h by ML/T, which is mv, or momentum.

Furthermore, the basic quantum relationship between energy and frequency $E = hv$ is assumed to hold for matter waves.

The AP exam does not require any application of de Broglie waves beyond being able to calculate the wavelength of a particle, but the idea of matter waves proved to be the central concept for the subsequent development of quantum mechanics. Eq. (12.11) does not say what a matter wave is; it does not say what is oscillating, be it an electric field or magnetic field or something else. (The answer is something else.) The theory created by Erwin Schrodinger and Werner Heisenberg to explain the behavior of matter waves became modern quantum mechanics.

Example 12.6: Davisson-Germer experiment. If matter behaves like waves, it should undergo diffraction and interference like the other waves we have discussed (Section 11.11).

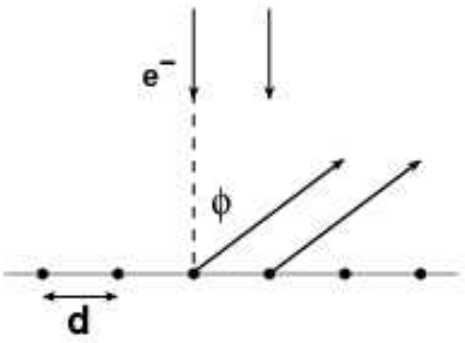


Figure 12.9

In 1927 Clinton Davisson and Lester Germer performed an experiment, which boiled down to its essentials, amounted to shining an electron beam of 54 eV electrons onto a crystal of nickel. The crystal acts like a **diffraction grating**, a screen with numerous narrow slits cut into it and which diffracts light much like the two-slit interferometer. In this case the role of the slits was played by the crystal atoms themselves, separated by a distance $d = 0.215$ nm. (This was known from previous experiments utilizing diffraction of x-rays.) Figure 12.9 shows the top row of such a nickel crystal. If two electron rays are diffracted off adjacent atoms, then the condition for constructive interference is the same as for the two-slit experiment:

$$d \sin \phi = m\lambda$$

(**Exercise:** From the illustration, convince yourself of this. In an emergency see Figure 11.24)

Davisson and Germer found a strong interference maximum for $m = 1$ at $\phi = 50°$, implying that $\lambda = 0.165$ nm. If the electron energy is 54 eV and we assume Newtonian physics, then the electron's kinetic energy is $1/2\, m_e v^2 = 54 \times 1.6 \times 10^{-19}$ J, where m_e is the electron mass. Solving for v gives us $v = 4.358 \times 10^6$. Thus, $p = m_e v = 3.965 \times 10^{-24}$ kg m/s. From Eq. (12.11) we find $\lambda = 0.167$ nm, in good agreement with the previous number. (One should point out the electrons are moving fast enough that special relativity should really be taken into account, which improves the result.)

10. Scientific History

The SAT subject test expects you to have a basic knowledge of the history of science. Scientific history, however, is far more complex than is presented in textbooks and what is stated in most undergraduate texts is, at best, a gross oversimplification, or at worst, grossly incorrect. Here is an oversimplified chronology and list of important figures, along with the discoveries associated with them. It more or less agrees with what is commonly found in texts. The most important thing to remember for an exam is that classical mechanics was developed in the seventeenth century, electromagnetism and thermodynamics in the nineteenth century, and modern physics in the late nineteenth and twentieth centuries.

Classical mechanics:

Nicholaus Copernicus (1473-1543)—Created a heliocentric model that placed the sun, not the earth, at the center of the solar system. The Greek Aristarchus made a similar proposal in the third century B.C.

Galileo Galilei (1564-1642)—Discovered the moons of Jupiter, as well as the kinematic relationships that lie at the basis of elementary physics. Proposed the law of inertia.

Johannes Kepler (1571-1630)—Discovered by trial and error the three laws of planetary motion that bear his name and led Isaac Newton to develop mechanics.

Isaac Newton (1643-1727)—The architect of mechanics. His three laws of motion form the basis of all classical physics. Believed that light was composed of particles.

Optics:

Christiaan Huygens (1629-1695)—Discovered the phenomenon of polarization. An early advocate of the idea that light is composed of waves, not particles. Also invented the pendulum clock.

Thomas Young (1773-1829)—Leading proponent of the theory that light is composed of waves. Universally cited as having proved this by the famous two-slit experiment. However, Young never mentions the experiment in his Royal Society lectures. In his single account in a Royal Institution lecture he never claims to have performed it.

Electricity and Magnetism:

Hans Ørsted (Oersted) (1777-1851)—On April 21, 1820 noticed that an electric current deflected a compass needle, thus proving that electricity and magnetism were inseparable phenomena and inaugurating the science of electromagnetism.

André-Marie Ampère (1775-1836)—Early pioneer in electricity and magnetism. Established the force between two current-carrying wires. He never postulated "Ampère's law," which gives the magnetic field around a wire. It was not until the 20th century that the law even became known as Ampère's. Maxwell may have been the first to propose it. The unit of current is named in Ampère's honor.

Michael Faraday (1791-1867)—Discovered electromagnetic inductance independently of Joseph Henry, discovered the principles of capacitance and electrolysis, which provided original definition of the amp. Unit of capacitance named in his honor.

Joseph Henry (1797-1878)—Made numerous important discoveries. Invented the first motor, created the world's strongest electromagnets, discovered electromagnetic induction independently of Michael Faraday. Invented the relay, which became the basis of Samuel Morse's telegraph. First Secretary of the Smithsonian Institution. Unit of inductance named in his honor.

James Clerk Maxwell (1831-1879)—Unified all previous work on electricity and magnetism into a single theory. Considered the creator of the modern theory of electromagnetism. Predicted the existence of electromagnetic waves.

Heinrich Hertz (1857-1894)—Generally credited with the discovery of electromagnetic waves, although Henry appears to have anticipated him, forty years earlier. Also discovered the photoelectric effect, assisted by Wilhelm Hallwachs (1859-1922) and which was independently discovered by Augusto Righi (1850-1920). Unit of frequency named in his honor.

Thermodynamics:

Robert Boyle (1627-1691)—Often regarded as the first modern chemist. Published Boyle's law relating pressure and volume, discovered by Richard Towneley and Henry Power. Also discovered independently by Edme Mariotte (1620-1684).

Daniel Bernoulli (1700-1782)—A century before atoms were accepted, he explained Boyle's law by assuming that gases were composed of invisible particles colliding with container. Also formulated Bernoulli's principle in fluid mechanics.

Sadi Carnot (1796-1832)—Idealized steam engines and invented the thermodynamic cycle. In doing so, made the first statement of the second law of thermodynamics.

Julius von Meyer (1814-1878)—Made first statement of conservation of energy. Also, along with Joule, discovered the equivalence of heat and energy.

James Prescott Joule (1818-1889)—With Meyer made the first argument for the mechanical equivalent of heat—that heat and energy were the same thing.

Rudolph Clausius (1822-1888)—Introduced the concept of entropy, made first clear statement of second law of thermodynamics.

William Thomson, Lord Kelvin (1824-1907)—Proposed the absolute temperature scale, put Carnot's work on a firmer footing by declaring one version of the second law of thermodynamics.

Radiation and radioactivity:

Claude Felix Abel Niepce de Saint Victor (1805-1870) —Discovered the darkening of photographic papers by uranium nitrate and published six memoirs on the subject beginning in 1856. Discovery debated and forgotten by contemporaries, repeated by Henri Becquerel in 1896.

Wilhelm Röntgen (1845-1923)—Discovered x-rays in 1895.

Henri Becquerel (1852-1908)—Discovered radioactivity due to uranium nitrate in 1896. Some evidence that he knew of Abel Niepce de St. Victor's earlier work.

Marie and Pierre Curie (1867-1934; 1859-1906)—Early investigators of radioactivity. Discovered radium and polonium.

Relativity, quantum, atomic and nuclear physics:

Henri Poincaré (1851-1912)—Provided crucial ideas that were incorporated into the special theory of relativity, for example that there is no such thing as absolute space or absolute time.

J.J. Thomson (1856-1940)—Made numerous contributions to physics. Known for discovering the electron in 1897, although Emil Weichert (1861-1928) and Walter Kauffmann (1871-1947) published similar results slightly earlier.

Max Planck (1858-1947)—Explained the blackbody spectrum in 1900 by assuming that energy comes in discrete units. In doing so he created quantum mechanics, which he spent years trying to disprove.

Ernest Rutherford (1871-1937)—Made numerous contributions to atomic and nuclear physics. Discovered alpha, beta and gamma radiation, radioactive decay. Proposed first successful model of the atom, based on famous "Rutherford experiment," carried out by Hans Geiger (1882-1945) and Ernest Marsden (1889-1970).

Albert Einstein (1879-1955)—Created the special and general theories of relativity, made fundamental contributions to quantum mechanics, in particular explained the photoelectric effect.

John William Nicholson (1881-1955)—First to propose that the electron's orbital angular momentum in an atom should be quantized. However, his own model of the atom was incorrect. Bohr openly acknowledged receiving the suggestion from Nicholson's papers.

Niels Bohr (1885-1962)—Explained the spectrum of atomic hydrogen by assuming that the angular momentum of the electrons was quantized.

Louis de Broglie (1892-1987)—Postulated the existence of matter waves. In doing so paved the way for modern quantum mechanics.

Erwin Schrödinger (1887-1961)—One of the creators of modern quantum mechanics. Proposed the "Schrödinger equation," the equation that underlies all quantum mechanics; equivalent to $F = ma$ for quantum systems.

Werner Heisenberg (1901-1976)—Created modern quantum mechanics independently of Schrödinger, via a different formalism. Also known for the "Heisenberg uncertainty principle," which states that it is impossible to simultaneously know the position and momentum of a particle with complete accuracy.

MODERN PHYSICS PROBLEMS

Problems 1-10 are at the subject-test level. The remaining problems are at the AP level.

1. The nucleus of an ordinary hydrogen atom consists of

   (A) a neutron.
   (B) a proton.
   (C) a proton and a neutron.
   (D) a proton and two neutrons.
   (E) two protons and two neutrons.

2. Of the following particles, which is the most massive?

   (A) photon
   (B) neutrino
   (C) neutron
   (D) proton
   (E) electron

3. Of the following scientists, who did not contribute to the development of quantum mechanics?

   (A) Niels Bohr
   (B) Max Planck
   (C) Albert Einstein
   (D) James Clerk Maxwell
   (E) John Nicholson

4. Two isotopes of an element ordinarily

   (A) have the same atomic number.
   (B) have the same atomic mass.
   (C) contain the same number of nucleons in the nucleus.
   (D) contain the same number of electrons.
   (E) (A) and (D)

5. Which of the following statements is false? When an atom is ionized,

   i. an electron is removed from the atom; ii, an electron is added to the atom; iii. the atom becomes charged; iv. the atom may be accelerated by an electric field; v. the energy of the electron is raised.

   (A) i only
   (B) ii only
   (C) i, ii, iv
   (D) ii, iii, iv
   (E) ii, iii, iv, v

6. The half-life of the hydrogen isotope tritium is about 12 years. After a certain amount of time a fraction 31/32 of the atoms in the original sample has decayed. The time is most nearly equal to

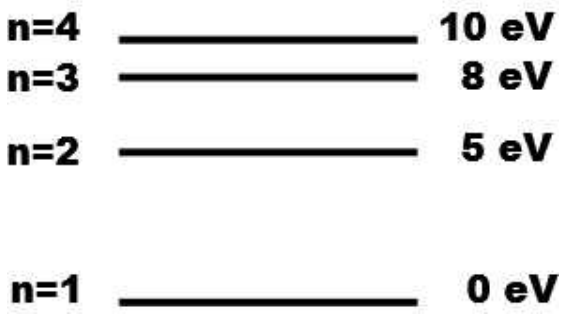
   (A) 12 years
   (B) 24 years
   (C) 36 years
   (D) 48 years
   (E) 60 years

7. A hydrogen atom in its ground state absorbs a photon. In the process an electron is excited to a higher state. The atom will then most likely

   (A) emit a photon.
   (B) remain in the excited state.
   (C) become ionized.
   (D) split in half.
   (E) collide with another atom.

8. Some energy levels in an imaginary atom are shown in the figure below.

If the atom absorbs a photon and makes a transition from a lower energy to a higher one, which of the following represent possible energies of the absorbed photon?



(A) 4 eV or 6 eV
(B) 2 eV or 5 eV
(C) 3 eV or 6 eV
(D) 4 eV or 10 eV
(E) 3 eV or 10 eV

9. Which of the following statements about the photon is true?

i. The photon has a wavelength; ii. The photon has a mass; iii. The photon can undergo interference; iv. The photon is charged; v. The photon has a momentum.

(A) i only
(B) i, ii only
(C) i, iii, v only
(D) i, ii, iii, iv only
(E) iii, v only

10. An electron in the Bohr atom has an energy level determined by the radius of the orbit around the nucleus. The lowest energy state is given by the Bohr radius, which is roughly 10^{-10} m. The radius of a hydrogen nucleus is about 10^{-14} m. If the nucleus were the size of a basketball, the size of the atom would be nearest to

(A) the size of a small town.
(B) the size of a supermarket.
(C) the size of a basketball court.
(D) the size of a continent.
(E) the size of the earth.

11. Who of the following did not pioneer the study of radioactivity?

(A) Marie Curie
(B) Pierre Curie
(C) Albert Einstein
(D) Felix Abel Niepce de Saint Victor
(E) Henri Becquerel

12. Which statements are true about Max Planck's quantum postulate? (More than one answer permitted.)

(A) The minimum energy of radiation emitted by a system can come only in discrete, or discontinuous, values.
(B) The energy radiated by a system can come only in integral multiples of Planck's constant.
(C) The energy radiated by a system can come only in integral multiples of Planck's constant multiplied by the frequency of the emitted light.
(D) The energy levels of a system are equally spaced by intervals of Planck's constant multiplied by the frequency of light.
(E) The energy radiated from a system must be in the form of photons.

13. An atomic nucleus is induced to break into two pieces, during which process energy is released. One can say with certainty that

    (A) the original atomic mass was greater than that of iron.
    (B) the fragments will each have an atomic mass less than that of iron.
    (C) the masses of the fragments will add up to be less than the mass of the original nucleus.
    (D) (A) and (C)
    (E) None of the above

14. A photon of energy 1.75 eV has a wavelength of most nearly

    (A) 710 nm
    (B) 700 nm
    (C) 650 nm
    (D) 600 nm
    (E) 550 nm

15. A laser used in retinal surgery emits photons of wavelength 632.8 nm in pulses of 0.5 W that last 2×10^{-2} s. The number of photons in each pulse is nearest to

    (A) 10^{10}
    (B) 10^{15}
    (C) 10^{20}
    (D) 10^{25}
    (E) 10^{30}

16. A naval destroyer has a mass of two million kilograms. To raise it fifty meters above the water requires the energy equivalent that is closest to the mass of

    (A) a virus, mass of about 10^{-18} kg
    (B) a bacterium, mass of about 10^{-15} kg
    (C) a human cell, mass of about 10^{-12} kg
    (D) a fruit fly, mass of about 10^{-7} kg
    (E) a mosquito, mass about 10^{-6} kg

17. Which of the following statements is true? The existence of the de Broglie wavelength λ_{dB} implies

    (A) that matter particles should undergo interference.
    (B) that matter waves travel at the speed of light.
    (C) that the frequency of matter waves is c/λ_{dB}, where c is the speed of the particle.
    (D) that matter waves are given off by accelerating charges.
    (E) that matter waves are polarized.

18. A virus has a mass of about 10^{-18} kg. The de Broglie wavelength of a virus being blown in the wind at 10 m/s is nearest to

    (A) one billionth the size of a hydrogen atom.
    (B) one millionth the size of a hydrogen atom.
    (C) one thousandth the size of a hydrogen atom.
    (D) the size of a hydrogen atom.
    (E) one thousand times the size of a hydrogen atom.

19. A boron atom bombarded by a slow neutron can undergo the following reaction:

$$\mathrm{n} + {}^{10}\mathrm{B} \rightarrow {}^{4}\mathrm{He} + {}^{7}\mathrm{Li} + \gamma.$$

The reaction products receive 2.31 MeV of kinetic energy. This kinetic energy should be computed as

(A) equal to the kinetic energy of the neutron.
(B) the total energy of the neutron.
(C) the energy deficit of the reaction.
(D) the energy deficit of the reaction minus the photon energy.
(E) the energy deficit of the reaction minus the kinetic energy of the neutron.

20. In the Bohr model of the atom, which of the following statements are false? (More than one response permitted.)

(A) The energy levels are spaced apart by hv.
(B) The angular momentum of the electron is quantized.
(C) The least amount of energy that can be absorbed by the atom is between the $n = 2$ state and the $n = 3$ state.
(D) The energy of a photon emitted in the Balmer series is proportional to $(\frac{1}{4} - \frac{1}{n^2})$ where n is the principal quantum number of the initial energy level.
(E) The energy levels characterized by the quantum number n are stable.

21. Measurements of the spectrum of a certain atom show that one sequence of spectral lines has the identical pattern of the Balmer series except that all the frequencies are exactly four times as high. The most likely explanation for the observation is that

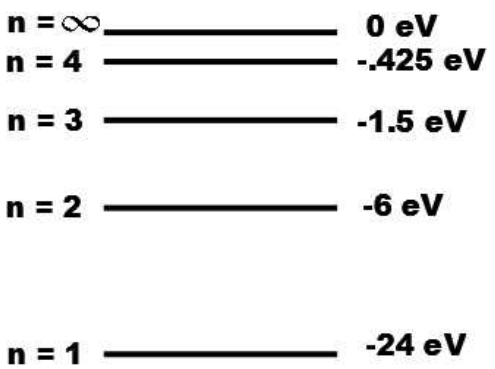
(A) one is observing a series of lines in hydrogen that have the ground state $n = 1$ instead of $n = 2$.
(B) the atom is helium, with atomic number two.
(C) the atom is beryllium with atomic number four.
(D) the atom is deuterium with atomic mass two.
(E) the atom is hydrogen but excited to energy levels that are four times higher than in the Balmer series.

22. The fusion reaction ${}^{2}_{1}\mathrm{H} + {}^{3}_{1}\mathrm{H} \rightarrow {}^{4}_{2}\mathrm{He} + \mathrm{n}$ releases about 17.5 MeV. The available energy is 14.1 MeV given to the neutron. The world's energy consumption in 2013 has been estimated as 4×10^{20} J. About one out of every 6420 hydrogen atoms in the earth's oceans is deuterium. The number of kilograms of water that would be enough to supply the world's annual energy needs is most nearly

(A) 10^{8} kg
(B) 10^{9} kg
(C) 10^{10} kg
(D) 10^{11} kg
(E) 10^{12} kg

Free-response problems

23. Illustrated below is an imaginary atom with several energy levels drawn.



a) Draw on the diagram all the allowed transitions when an electron falls from the $n = 4$ or $n = 3$ states to a lower energy.

b) What is the shortest possible wavelength of light emitted during such transitions?

c) What is the longest possible wavelength of light emitted during such transitions?

d) An electron is in the $n = 2$ state. A photon of energy 3 eV is incident on the atom. What effect does it have on the electron?

e) An electron is in the $n = 2$ state. A photon of energy 8 eV is incident on the atom. What is the effect on the electron?

f) What is the ionization potential of the atom?

24. A hypothetical atom has energy levels at -12 eV, -8 eV, -3 eV, -1 eV.

a) Draw the energy levels of the atom. Label the levels with the principal quantum number.

b) An electron with velocity $v = 1.326 \times 10^6$ m/s is incident on the atom. What is the de Broglie wavelength of the electron? Ignore any relativistic effects.

c) Can the electron excite an electron in any of the energy levels to a higher state? If so, which are the two levels involved?

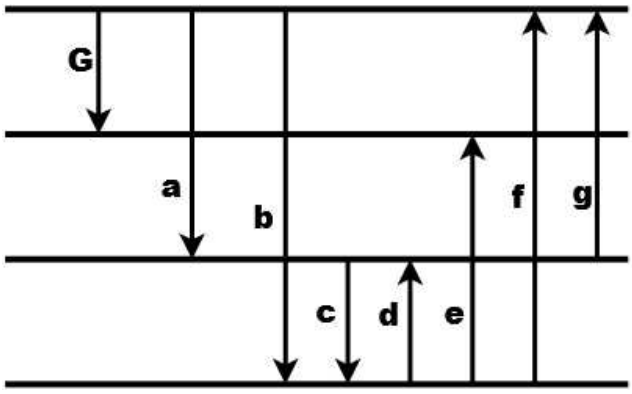
d) An electron decays from the -3 eV state to the -8 eV state and emits a photon. What is its wavelength? What part of the electromagnetic spectrum is it in?

25. A gas consists of atoms that are known to have a ground state at -6 eV. When a white light source is shone on the gas, a spectrometer sensitive to the visible spectrum determines that light is absorbed at a wavelength of 414 nm. When the light source is switched off, the spectrometer determines that light is emitted by the gas at 414 and 621 nm.

a) Draw the energy levels of the atom that the observations suggest and draw in the possible transitions. Draw in at least one other energy level that the atom could have.

b) Why aren't emission lines of other frequencies observed?

26. A photon of wavelength 200 nm ejects an electron from a metal. The de Broglie wavelength of the ejected electron is 1.17 nm.

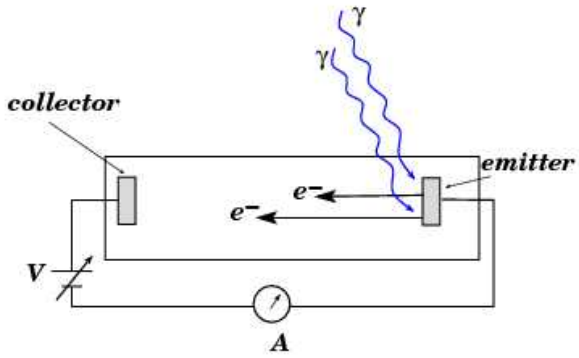


a) What is the kinetic energy of the electron in eV?

b) What is the work function of the metal?

c) The 200 nm photon was created in an atomic transition. The transition labeled G to the right created a 600 nm photon. Which of the other indicated transitions could have created the 200 nm photon?

Questions 27 - 28 refer to the following figure.



27. A light source of variable frequency and intensity is directed onto an unknown piece of metal (emitter). A variable voltage source (V) is initially set to zero. At a threshold frequency of 8.82×10^{14} Hz, the emitter begins to emit photoelectrons, which cause a current that is measured by the ammeter (A). The frequency of the light is further increased and the energy of the electrons recorded.

a) What is the wavelength of the incident photons at the threshold? In what part of the electromagnetic spectrum is this wavelength?

b) What is the work function of the metal?

c) The voltage on variable supply (V) is increased from zero until the electrical current drops to zero. The voltage at which this occurs is 1.2 V. Assuming Newtonian physics, what is the maximum velocity of the electrons?

d) With the electrical current at zero, the intensity of the light (watts per square meter) is increased. This will

cause the electrical current to reappear____ have no effect____
Justify your response.

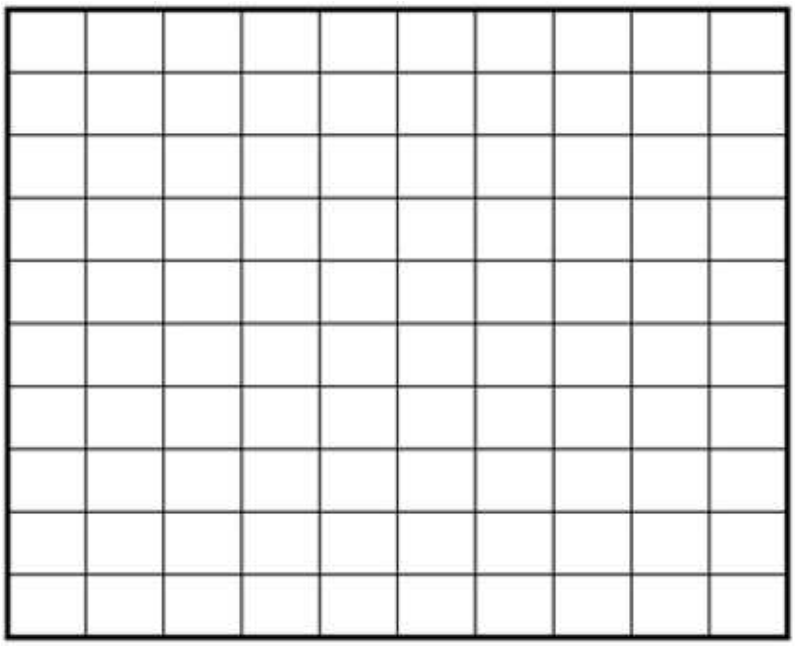
e) With the electrical current at zero and the light intensity at its initial value, the frequency of light is increased. This will

cause the electrical current to reappear____ have no effect____
Justify your response.

28. You shine a light source of variable wavelength on the same apparatus as above, which contains an unknown metal for the emitter. As you decrease the wavelength, nothing happens until you reach a wavelength of 540 nm, when you begin to record a small current. You adjust the voltmeter to find that the current stops at a voltage of 0.02 V. You continue to decrease the light wavelength, recording the following data:

Wavelength	Stopping Voltage
540 nm	0.02 V
500 nm	0.14 V
460 nm	0.39 V
420 nm	0.70 V

You are given a piece of graph paper.



Plot a graph from which you can determine Planck's constant.

29. In 1923 Arthur Compton observed that x-rays scattered off free electrons were shifted in wavelength by an amount that could be explained by assuming that the x-rays and electrons obeyed the relativistic relationships for energy and momentum. Compton used x-rays (γ) with a wavelength of approximately 7.30×10^{-11} m. Assume that such a photon is incident on a stationary electron, as shown below.

For purposes of illustration, the photon is reflected off the electron and its wavelength is observed to shift by a magnitude $|\Delta\lambda| = |\lambda_o - \lambda_f| = h/m_e c$, where λ_o and λ_f are the initial and final wavelengths, respectively.

a) What is the energy of the incoming photon in eV? in joules?

b) Does the wavelength of the photon increase or decrease? Explain your reasoning.

c) What is the momentum acquired by the electron?

30. A plutonium-239 nucleus at rest spontaneously undergoes fission via the following reaction:

$$^{239}_{94}\text{Pu} \rightarrow {}^{112}_{46}\text{Pd} + {}^{124}_{48}\text{Cd} + 3\text{n}.$$

The mass of plutonium-239 in atomic mass units is $^{239}\text{Pu} = 239.0521$. The mass of palladium-112 is 111.9073. The mass of cadmium-124 is 123.9176. The mass of a neutron is 1.0087.

a) What is the Q value of the reaction?

b) If we ignore the neutrons, what should be the speeds of the palladium and cadmium nuclei? (Assume Newtonian physics.)

c) If the palladium and cadmium nuclei can be treated as spheres, and are initially far apart, how close can they get to each other at these velocities?

d) ^{112}Pd decays via beta emission to ^{112}At with a half-life of 21 hrs. If a sample contains 97% argon, how long has the palladium been decaying?

31. Every subatomic particle has an antiparticle counterpart. An antiparticle is a particle with the same mass as the given particle, but with opposite electric charge. The **positron** (e^+) is the electron's antiparticle counterpart; it is identical to the electron except that it carries a positive charge. In certain circumstances a particle and its antiparticle can be created together; such a process is called **pair production**. Furthermore, the energy equivalent of the mass of a particle at rest is termed the **rest energy**.

a) If the electron's mass is 9.1×10^{-31} kg, what is its rest energy in eV?

b) What is the minimum energy a photon must have in order to decay into an electron-positron pair? (Express your answer in eV.) What is its frequency, wavelength and momentum?

c) Is it actually possible for a free photon to spontaneously decay into an electron-positron pair? Justify your answer.

32. Meteorites created in the early solar system contained aluminum-26, which is a radioactive isotope of aluminum with a half-life of 7.2×10^5 yrs. Aluminum-26 decays first into an excited state of magnesium-26 via the reaction ${}^{26}_{13}\text{Al} \rightarrow {}^{26}_{12}\text{Mg}^* + \text{e}^+$, where the e^+ has energy 2.99 MeV. (The e^+ is a positron; see previous problem. The asterisk (*) indicates "excited.") The ${}^{26}_{12}\text{Mg}^*$ then decays into the stable isotope magnesium-26 via the reaction ${}^{26}_{12}\text{Mg}^* \rightarrow {}^{26}_{12}\text{Mg} + \gamma$. The γ has energy 1.8 MeV.

a) If you were asked to calculate the de Broglie wavelength of the positron, would it be permissible to use Newtonian physics? Justify your answer.

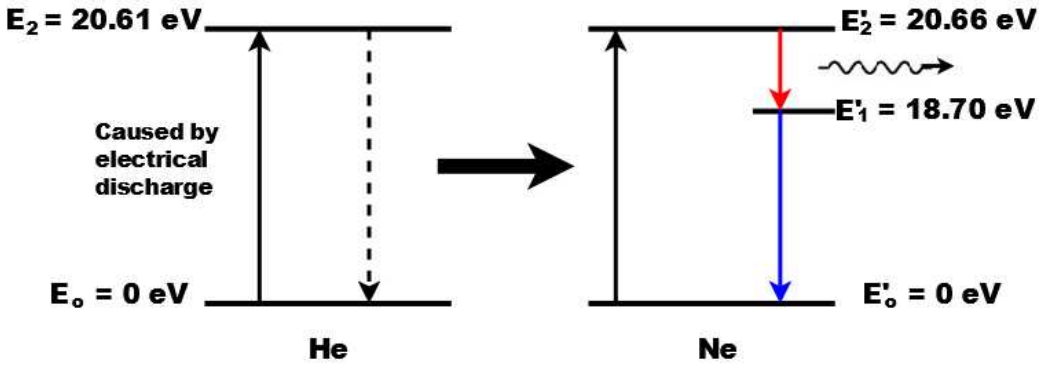
b) What is the wavelength of the photon emitted when the excited magnesium-26 decays into its ground state? What is its momentum?

c) ${}^{26}_{12}\text{Mg}$ has an atomic mass of 25.9826 u. What is the speed of the recoiling nucleus when the photon is emitted?

d) What is the nucleus' kinetic energy in electron volts?

e) Precise measurements indicate that for a certain meteorite *A* the present ratio ${}^{26}\text{Mg}/{}^{27}\text{Al} = 5 \times 10^{-5}$, where ${}^{27}\text{Al}$ is the common, stable isotope of aluminum. For a meteorite *B* the ratio is ${}^{26}\text{Mg}/{}^{27}\text{Al} = 1.55 \times 10^{-7}$. Assuming that the different ratios are due to the difference in the meteorites' times of creation, how much older is meteorite *B* than *A*?

33. We end the problem sequence in this book with the following question, which appeared on the 2003 AP Physics B exam. It concerns the operation of a helium-neon laser and wins our nomination for the most confusing AP-exam problem (or solution) ever written.



The introduction to the problem reads:

Energy-level diagrams for atoms that comprise a helium-neon laser are given above. As indicated on the left, the helium atom is excited by an electrical discharge and an electron jumps from energy level E_o to energy level E_2. The helium atom (atomic mass 4) then collides inelastically with a neon atom (atomic mass 20), and the helium atom falls to the ground state, giving the neon atom enough energy to raise an electron from E_o' to energy level E_2'. The laser emits light when an electron in the neon atom falls from energy level E_2' to energy level E_1'.

The large horizontal arrow in the original diagram is labeled "inelastic collision." The question then asks:

Calculate the minimum speed the helium atom must have in order to raise the neon electron from E_0' to $E_2{}'$.

About the Author

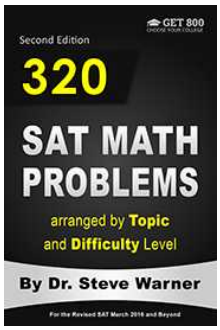
Tony Rothman is a theoretical cosmologist who holds a PhD in physics from the University of Texas, Austin. In 2016 he joined the faculty of NYU's Tandon School of Engineering, having previously taught at Princeton and Harvard Universities. To date Rothman has authored approximately 60 scientific papers as well as 11 previous books, both fiction and nonfiction. The latest of these are two novels, *Firebird*, about a race for nuclear fusion; and *The Course of Fortune*, a historical novel about the Great Siege of Malta in 1565. His book *Sacred Mathematics, Japanese Temple Geometry*, with Hidetoshi Fukagawa, won the 2008 American Association of Publishers Award for Professional and Scholarly Excellence in mathematics. He has won numerous other writing awards and has been nominated for the Pulitzer Prize.

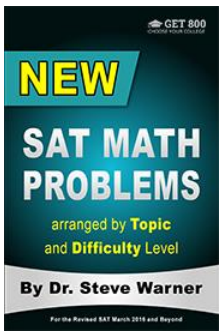
About the Editor

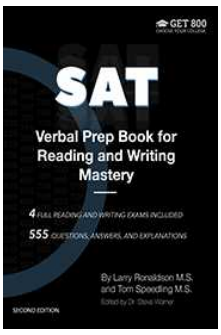
Dr. Steve Warner earned his Ph.D. at Rutgers University in Mathematics, and he currently works as an Associate Professor at Hofstra University. Dr. Warner has more than 20 years of experience in general math tutoring and more than 15 years of experience in SAT math tutoring. He has tutored students both individually and in group settings and has published several math prep books for the SAT, ACT and AP Calculus exams.

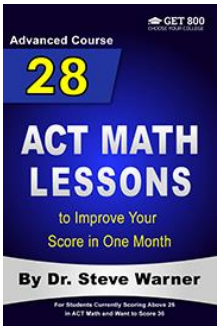
CONNECT WITH DR. STEVE WARNER

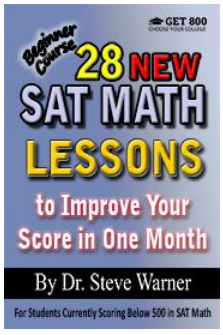
BOOKS FROM THE GET 800 COLLECTION

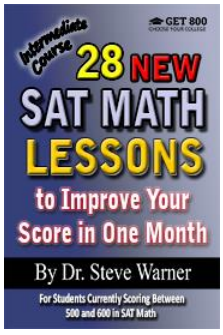


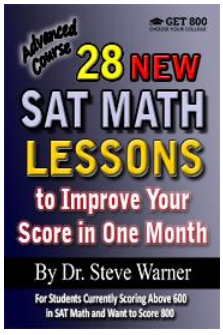


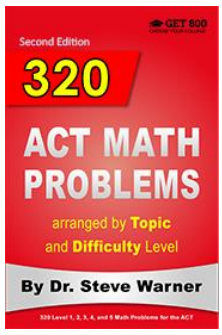

GET 800
SAT
Verbal Prep Book for Reading and Writing Mastery
4 FULL READING AND WRITING EXAMS INCLUDED
555 QUESTIONS, ANSWERS, AND EXPLANATIONS
By Larry Ronaldson M.S. and Tom Speedling M.S.

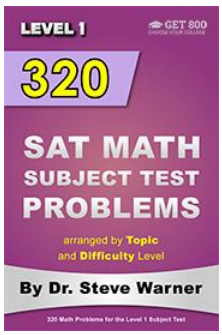


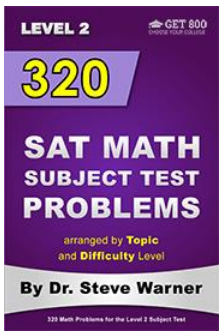

GET 800
Beginner Course
28 NEW
SAT MATH LESSONS
to Improve Your Score in One Month
By Dr. Steve Warner
For Students Currently Scoring Below 500 in SAT Math

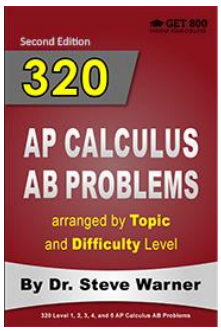

GET 800
Intermediate Course
28 NEW
SAT MATH LESSONS
to Improve Your Score in One Month
By Dr. Steve Warner
For Students Currently Scoring Between 500 and 600 in SAT Math

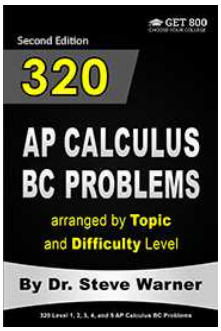

GET 800
Advanced Course
28 NEW
SAT MATH LESSONS
to Improve Your Score in One Month
By Dr. Steve Warner
For Students Currently Scoring Above 600 in SAT Math and Want to Score 800

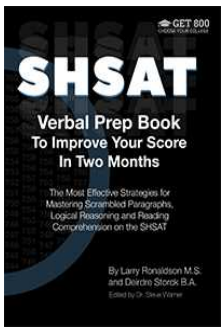


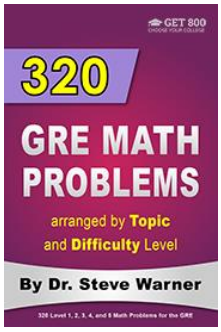


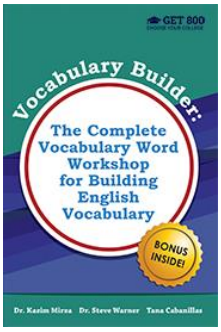


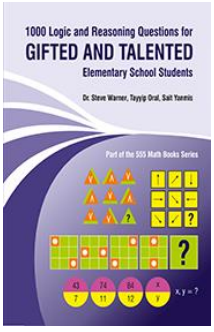





GET 800
SHSAT
Verbal Prep Book To Improve Your Score In Two Months
The Most Effective Strategies for Mastering Scrambled Paragraphs, Logical Reasoning and Reading Comprehension on the SHSAT
By Larry Ronaldson M.S. and Deirdre Storck B.A.







GET 800
CHOOSE YOUR COLLEGE

Physics Mastery, a complete course written by a Ph.D. in physics, emphasizes the topics and question types that appear on the SAT physics subject test, as well as on the AP physics 1 and AP physics 2 exams. Each chapter has been carefully constructed to provide you with clever and efficient ways of solving problems to ensure that you spend less time on each question.

The book consists of 12 comprehensive yet concise review chapters that cover all of the concepts and strategies needed to take you to a perfect score on the SAT subject test and AP physics exams. Of equal importance are the 400 problems and 100 worked examples that give you the practice, math techniques and physical intuition you need to be prepared for the College Board exams, as well as to understand a freshman physics course.

Each problem is followed by at least one complete solution (often more), helpful remarks to increase comprehension, tricks of the trade that allow you to bypass tedious calculations, and reality checks that ensure you have the right answer.

Here's to your success on the College Board exams, in college, and in life.